Title

Unraveling the functional role of the orphan solute carrier, SLC22A24 in the transport of steroid conjugates through metabolomic and genome-wide association studies.

Journal

PLoS genetics, 15(9)

ISSN

1553-7390

Authors

Yee, Sook Wah
Stecula, Adrian
Chien, Huan-Chieh
et al.

Publication Date

2019-09-01

DOI

10.1371/journal.pgen.1008208

Peer reviewed

PLOS GENETICS

RESEARCH ARTICLE

Unraveling the functional role of the orphan solute carrier, SLC22A24 in the transport of steroid conjugates through metabolomic and genome-wide association studies

Sook Wah Yee[1☯], Adrian Stecula[1☯], Huan-Chieh Chien[1], Ling Zou[1], Elena V. Feofanova[2], Marjolein van Borselen[1], Kit Wun Kathy Cheung[1], Noha A. Yousri[3,4], Karsten Suhre[5], Jason M. Kinchen[6], Eric Boerwinkle[2,7], Roshanak Irannejad[8], Bing Yu[2], Kathleen M. Giacomini[1,9]*

1 Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, California, United States of America, **2** Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas, United States of America, **3** Genetic Medicine, Weill Cornell Medicine-Qatar, Doha, Qatar, **4** Computer and Systems Engineering, Alexandria University, Alexandria, Egypt, **5** Physiology and Biophysics, Weill Cornell Medicine-Qatar, Doha, Qatar, **6** Metabolon, Inc, Durham, United States of America, **7** Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, United States of America, **8** The Cardiovascular Research Institute, University of California, San Francisco, California, United States of America, **9** Institute for Human Genetics, University of California San Francisco, California, United States of America

☯ These authors contributed equally to this work.
* kathy.giacomini@ucsf.edu

OPEN ACCESS

Citation: Yee SW, Stecula A, Chien H-C, Zou L, Feofanova EV, van Borselen M, et al. (2019) Unraveling the functional role of the orphan solute carrier, SLC22A24 in the transport of steroid conjugates through metabolomic and genome-wide association studies. PLoS Genet 15(9): e1008208. https://doi.org/10.1371/journal.pgen.1008208

Editor: William E. Evans, St. Jude Children's Research Hospital, UNITED STATES

Received: May 20, 2019

Accepted: July 17, 2019

Published: September 25, 2019

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: This work was supported by the National Institutes of Health [GM117163 and DK108722 to K.M.G. and S.W.Y., R00HL122508 to R.I.]. The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute,

Abstract

Variation in steroid hormone levels has wide implications for health and disease. The genes encoding the proteins involved in steroid disposition represent key determinants of interindividual variation in steroid levels and ultimately, their effects. Beginning with metabolomic data from genome-wide association studies (GWAS), we observed that genetic variants in the orphan transporter, SLC22A24 were significantly associated with levels of androsterone glucuronide and etiocholanolone glucuronide (sentinel SNPs p-value $<1x10^{-30}$). In cells over-expressing human or various mammalian orthologs of SLC22A24, we showed that steroid conjugates and bile acids were substrates of the transporter. Phylogenetic, genomic, and transcriptomic analyses suggested that SLC22A24 has a specialized role in the kidney and appears to function in the reabsorption of organic anions, and in particular, anionic steroids. Phenome-wide analysis showed that functional variants of SLC22A24 are associated with human disease such as cardiovascular diseases and acne, which have been linked to dysregulated steroid metabolism. Collectively, these functional genomic studies reveal a previously uncharacterized protein involved in steroid homeostasis, opening up new possibilities for SLC22A24 as a pharmacological target for regulating steroid levels.

National Institutes of Health, Department of Health and Human Services (contract numbers HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I and HHSN268201700005I). The authors thank the staff and participants of the ARIC study for their important contributions. Funding support for "Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium" was provided by the NIH through the American Recovery and Reinvestment Act of 2009 (ARRA) (5RC2HL102419). Metabolomics measurements were sponsored by the National Human Genome Research Institute (3U01HG004402-02S1). Sequencing was carried out at the Baylor College of Medicine Human Genome Sequencing Center and supported by the National Human Genome Research Institute grants U54 HG003273 and UM1 HG008898. K.S. and N.A. Y. were supported by the Biomedical Research Program funds at Weill Cornell Medical College in Qatar, a program funded by the Qatar Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: J.M.K. is employed by Metabolon Inc. Other authors have declared that no competing interests exist.

Author summary

Steroid hormones, ranging from sex steroids such as testosterone to glucocorticoids play key roles in human health and disease. Accordingly, the identification of the genes and proteins involved in their synthesis, disposition and elimination has been the subject of numerous genetic studies. We have been intrigued by recent studies demonstrating that genetic variants in or near a gene encoding SLC22A24 are strongly associated with steroid levels. SLC22A24 is an orphan transporter with no known ligands and no known biological functions. In this study, we use cellular and computational methods to show that SLC22A24 transports steroid conjugates, bile acids, and other dicarboxylic acids. Based on the direction of association of a common stop codon in SLC22A24 with lower levels of steroids, our studies suggest that the transporter functions to reabsorb steroid conjugates in the kidney, a surprising finding, given that conjugation pathways generally function to create polar molecules that are readily eliminated by the kidney. The absence of the transporter gene in many species and its presence in higher order primates both suggest that SLC22A24 plays a specialized role in steroid homeostasis. Overall, our studies indicate that SLC22A24 functions in the reabsorption of conjugated steroids in the kidney.

Introduction

Steroid hormones represent major signaling molecules and have been the subject of numerous investigations[1]. Precise regulation of steroid levels is important for human health and disease, which has stimulated many studies focused on the identification and characterization of the genes, proteins, and pathways involved in steroid metabolism and disposition. In addition, genome-wide association studies (GWAS) have been performed to identify the polymorphisms that associate with interindividual variation in the levels of steroids, and to discover additional genes that are involved in steroid disposition. In general, multiple GWAS have led to the identification of genes already known to be involved in steroid metabolism, binding and regulation [2–7]. However, one locus, *SLC22A24*, has puzzled researchers because it has no known role in steroid homeostasis [6, 8, 9]. *SLC22A24* encodes an orphan solute carrier (SLC) transporter which has no currently known ligands.

The human solute carrier (SLC) transporter superfamily includes approximately 400 proteins grouped into 52 families [10–13]. SLC22A24, an orphan transporter, belongs to the SLC22 family, which transports organic ions [14]. In the human genome, the SLC22 family includes 22 members and represents the third largest in the SLC transporterome, behind SLC25 and SLC35 families. The first human member of this family was cloned and characterized in 1997 [15, 16], and since then, members of the family were shown to play critical pharmacological roles in mediating the hepatic and renal elimination of a variety of endogenous compounds and drugs (e.g., *SLC22A1* encoding organic cation transporter 1 (OCT1), *SLC22A2* encoding OCT2, *SLC22A6* encoding organic anion transporter 1 (OAT1), and *SLC22A8* encoding OAT3)[17, 18]. Because of their important pharmacologic role, many members of the SLC22 family are recognized as important determinants of drug disposition and mediators of drug-drug interactions [17], and are frequently studied during drug development [12].

In spite of the increasing recognition as playing key roles in drug response and toxicity, and indeed in human physiology and pathophysiology, the SLC22 family still contains a large cluster of orphan transporters, including SLC22A10, SLC22A14, SLC22A15, SLC22A17, SLC22A18, SLC22A23, SLC22A24, SLC22A25, and SLC22A31 [19, 20]. In fact, close to 75

members of the human SLC superfamily remain classified as orphan transporters, i.e., they have no assigned ligands [10, 11, 21]. A number of factors appear to contribute to this paradox including a non-unified nomenclature and technical challenges involved in working with membrane proteins[11].

SLCs that have been characterized to date function in the transmembrane flux of solutes, which include inorganic ions, such as heavy metals, as well as organic compounds, such as neurotransmitters and amino acids. Therefore, the key step in understanding the physiological role of these proteins is identifying their endogenous substrates.

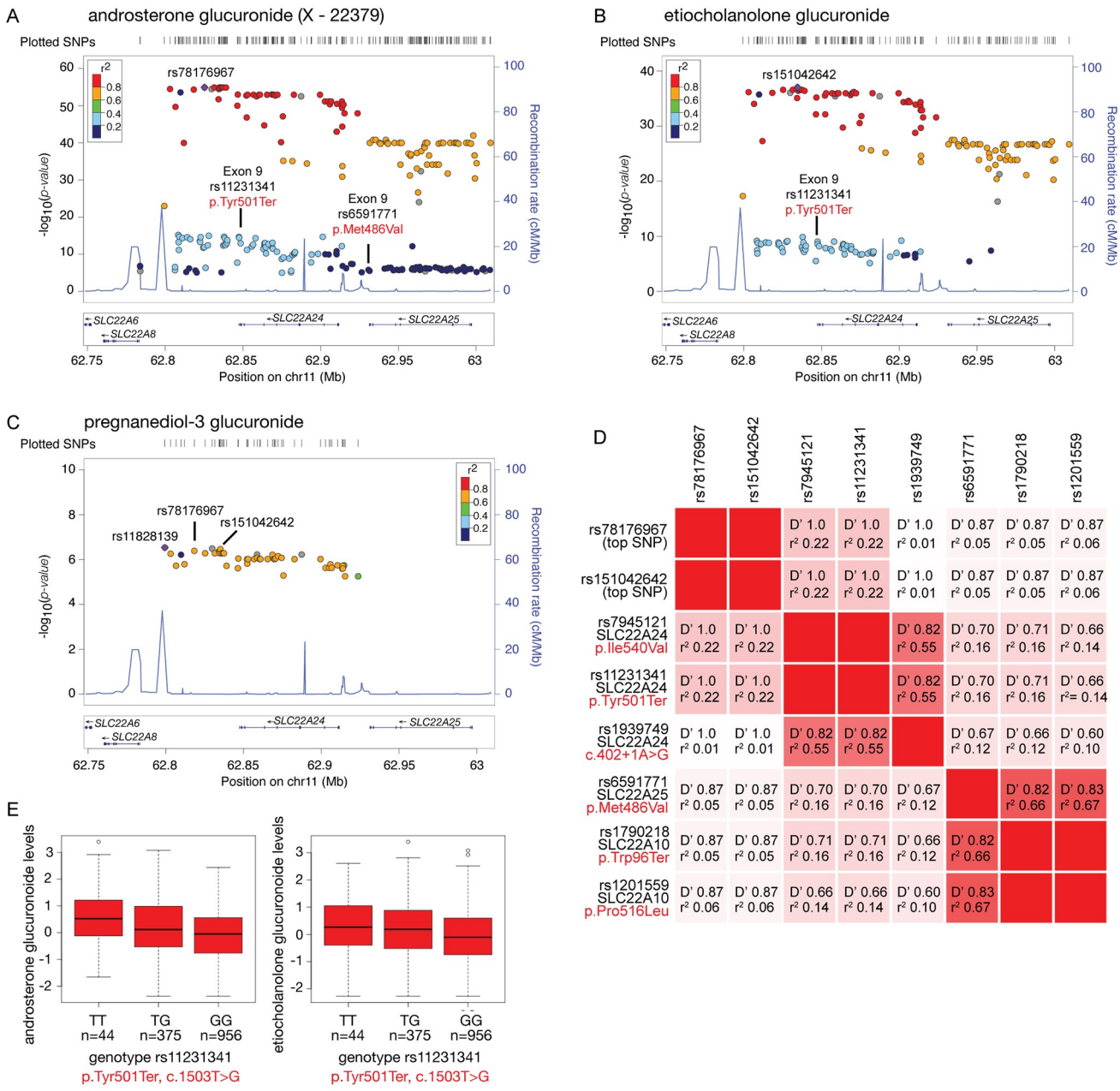
Over the years, various functional and molecular approaches have been employed to identify the endogenous ligands of SLC transporters. One robust approach for the identification of substrates has been the use of genome-wide association studies (GWAS) to identify single nucleotide polymorphisms (SNPs) in SLCs that associate with circulating levels of various metabolites. For example, the discovery of GLUT9 (encoded by *SLC2A9*) as a uric acid transporter stemmed from a GWAS focused on the identification of genetic variants associated with hyperuricemia and gout [22, 23]. Further, the characterization of SLC16A9 as a carnitine efflux transporter was motivated by a GWAS with metabolomics [24]. Also, tetradecanedioic acid and hexadecanedioic acid were discovered to be naturally occurring substrates of OATP1B1 (encoded by *SLCO1B1*) in a GWAS aimed at identifying metabolomic biomarkers for the transporter [25, 26].

In this study, we used a combination of computational and experimental approaches to uncover endogenous ligands and substrates of SLC22A24. These approaches included radiolabeled substrate and inhibition screens, metabolomic profiling of cells over-expressing SLC22A24, and sequence and structure-based analyses. To provide further support for the endogenous role of the human SLC22A24, we profiled its expression levels in various tissues and cell types. Phylogenetic analysis revealed that the transporter has orthologs in primates and a few other mammalian species, suggesting a specialized role of the transporter. Our work also implicates the transporter in the disposition of drugs and drug metabolites with a steroid scaffold. To our knowledge, this is the first study to identify endogenous and pharmacological substrates of SLC22A24. Further studies are warranted to understand the complete physiological and pharmacological roles of the transporter in human and primate biology.

Results

GWAS reveal strong associations of *SLC22A24* polymorphisms with steroid metabolite levels

Three independent GWAS reveal strong associations of SNPs in *SLC22A24* with steroid levels. In particular, the metabolites that reached genome-wide significance included progesterone ($p<10^{-11}$)[6], androsterone glucuronide (sentinel SNP $p<10^{-50}$)[8, 9], and etiocholanolone glucuronide (sentinel SNP $p<10^{-30}$)[8]. A locus zoom plot shows the significance pattern of SNPs in and around *SLC22A24* for the association with androsterone glucuronide (Fig 1A), etiocholanolone glucuronide (Fig 1B) and pregnanediol-3-glucuronide (Fig 1C). The SNPs in the vicinity of *SLC22A24* have the most significant associations with androsterone glucuronide and etiocholanolone glucuronide in GWAS [8, 9]. The top SNPs for progesterone (rs112295236) and the sentinel SNPs for androsterone glucuronide and etiocholanolone glucuronide (rs78176967, rs151042642)[6, 8, 9] are in the region between *SLC22A24* and *SLC22A8*. These SNPs are in strong linkage disequilibrium (LD), $r^2>0.8$, to SNPs within *SLC22A24* (Fig 1A and 1B). Interestingly, a nonsense variant in *SLC22A24* (rs11231341, Tyr501Ter, c.1503T>G), is strongly associated with the levels of androsterone glucuronide and etiocholanolone glucuronide ($p<5x10^{-8}$) (Fig 1A and 1B). Using LDLink[27], we identified the SNPs in



F

rs number	position (GRCh37)	allele frequencies (minus strand)	haplotypes			
rs78176967	chr11:62825317	C=0.945,A=0.055	C	C	A	C
rs151042642	chr11:62834557	G=0.945,A=0.055	G	G	A	G
rs11231341	chr11:62848487	G=0.792, T=0.208	G	T	T	G
rs1939749	chr11:62910849	G=0.825,A=0.175	G	A	G	A
		haplotype count	772	151	55	25
		haplotype frequency	0.7674	0.1501	0.0547	0.0249

Fig 1. Genetic associations of variants in the *SLC22A24* locus with steroid glucuronide levels. LocusZoom plot showing association results of top variants in the *SLC22A24* locus with (A) androsterone glucuronide (X-22379), (B) etiocholanolone glucuronide and (C) pregnanediol-3 glucuronide. The LD estimates are color coded as a heatmap from dark blue ($0<r^2<0.2$) to red ($0.8<r^2<1.0$). Recombination rate are indicated by the solid blue lines (recombination rate in cM/Mb from HapMap). The bottom panel shows the genes and their orientation. The associations were plotted using available data from Long T. et al. (http://www.hli-opendata.com/Metabolome/, S1 Table). Underlying data are also provided in S1 Data. (D) The linkage disequilibrium as calculated by D prime (D') and correlations (r^2) between the top SNPs rs781769767 and rs151042642 with missense and nonsense variants in *SLC22A24*, *SLC22A25* and *SLC22A10*. (E) Genetic association of *SLC22A24* p.Tyr501Ter with normalized residuals of androsterone glucuronide and etiocholanolone glucuronide in the ARIC study (N = 1375 Caucasians). The major allele of rs11231341 coding for p.Tyr501Ter is significantly associated

with lower serum levels of androsterone glucuronide and etiocholanolone glucuronide (see Table 1). Underlying data are provided in S1 Data. (F) Haplotype frequencies of the haplotypes observed for the list of query variants in European population from Phase 3 of the 1000 Genomes Project. This plot is created using LDhap https://ldlink.nci.nih.gov/?tab=ldhap[27].

https://doi.org/10.1371/journal.pgen.1008208.g001

linkage with rs78176967 and rs151042642, and found that among the missense, nonsense, or splice donor variants in *SLC22A24* (rs11231341, rs7945121, rs1939749), *SLC22A25* (rs6591771) and *SLC22A10* (rs1790218, rs1201559), the *SLC22A24* p.Tyr501Ter (rs11231341) has the strongest correlation to the top SNPs associated with the steroid glucuronides, with r^2 = 0.22, D' = 1.0 (Fig 1D). In addition to the published data, we used data from the Atherosclerosis Risk in Communities Study (ARIC) to confirm the direction of the association of the nonsense variant, *SLC22A24* p.Tyr501Ter, with plasma levels of the two metabolites. In particular, this independent cohort consists of 1,375 Caucasians and 573 African Americans with available genotype data and steroid metabolite levels. Primary analysis showed that the major allele (G-allele), encoding the nonsense variant SLC22A24 p.Tyr501Ter, is associated with lower plasma levels of the two steroid glucuronide conjugates, androsterone glucuronide and etiocholanolone glucuronide (Fig 1E, Table 1). Further, association analysis in ARIC study also confirmed weaker associations of other missense variants or nonsense variants in *SLC22A25* and *SLC22A10* (S2 Table).

We were puzzled by the observation that the sentinel SNPs, rs78176967 and rs151042642, exhibited stronger associations with steroid levels (p-values range from $1x10^{-55}$ to $9x10^{-11}$) than the nonsense SNP, rs11231341 (encoding p.Tyr501Ter, p-values range from 0.0014–$1x10^{-13}$) (Table 2). As the sentinel SNPs, rs78176967 and rs151042642, are intergenic variants with no known or predicted cis-regulatory effects, we suspected that they may tag true causal variants that are linked to them in addition to rs11231341. Searching for publicly available eQTL, we noted that the splice donor variant in *SLC22A24*, rs1939749, is the top SNP for associations with *SLC22A24* transcript levels in kidney eQTL databases[28–30] (S1 Fig). In particular, rs1939749 resides at the 3' junction of exon 1 and could affect splicing between exons 1 and 2. Incorrect splicing at the 3' junction of exon 1, i.e., with the minor allele variant,

Table 1. Association of SLC22A24 p.Tyr501Ter (rs11231341, c.1503T>G) with two steroid conjugates in three cohorts. Association of SLC22A24 p.Tyr501Ter (rs11231341, c.1503T>G) with steroid glucuronide levels. Major G-allele (coding for SLC22A24 p.Tyr501Ter) is associated with lower steroid glucuronide levels. The data from ARIC cohort are available through the established database of Genotypes and Phenotypes (dbGaP) application procedure for the ARIC study (dbGaP Study Accession: phs000280.v5.p1).

SNP (rsID)	A1	A2	Effect Allele	Effect Allele Freq	N	Beta	SE	P-value	Cohort	Metabolite	Source
rs11231341	T	G	G	0.83	1375	-0.39	0.051	**1.73E-13**	ARIC (European)	**Androsterone glucuronide (X-22379)**	This study
rs11231341	T	G	G	0.62	573	-0.20	0.064	0.0014	ARIC (African American)		This study
rs11231341	T	G	G	0.79	564	-0.30	0.067	5.43E-06	Middle Eastern		Using data computed from Yousri et al.[9]
rs11231341	T	G	G	0.81	1895	-0.36	0.053	**1.06E-13**	European (Long T)		Reference[8]
rs11231341	T	G	G	0.83	1375	-0.29	0.051	**3.14E-08**	ARIC (European)	**Etiocholanolone glucuronide**	This study
rs11231341	T	G	G	0.62	573	-0.21	0.064	0.0012	ARIC (African American)		This study
rs11231341	T	G	G	0.79	554	-0.34	0.072	3.41E-06	Middle Eastern		Using data computed from Yousri et al.[9]
rs11231341	T	G	G	0.81	1903	-0.26	0.053	**6.04E-09**	European (Long T)		Reference[8]

https://doi.org/10.1371/journal.pgen.1008208.t001

Table 2. Association of sentinel SNPs with two steroid conjugates in three cohorts. Association of sentinel SNPs with steroid glucuronide levels in two published studies and in the ARIC cohort. According to 1000 Genome Project, European population, these SNPs (rs78176967, rs61285056 and rs151042642) are in strong LD with one another (D' = 1, r^2>0.95). The A1 and A2 alleles are on the minus strand. The primary data are available through the established dbGaP application procedure for the ARIC study (dbGaP Study Accession: phs000280.v5.p1).

SNP (rsID)	A1	A2	Effect Allele	Effect Allele Freq	N	Beta	SE	P-value	Cohort	Metabolite	Source
rs78176967	C	A	A	0.063	1895	1.14	0.08	**1.17E-55**	European (Long T)	**Androsterone glucuronide (X-22379)**	Reference[8]
rs61285056	C	A	A	0.058	564	0.77	0.12	**9.11E-11**	Middle Eastern		Reference[9]
rs61285056	C	A	A	0.064	1456	1.17	0.077	**7.03E-53**	ARIC (European)		This study
rs151042642	G	A	A	0.063	1903	0.90	0.09	**9.13E-38**	European (Long T)	**Etiocholanolone glucuronide**	Reference[8]
rs61285056	C	A	A	0.061	1456	0.95	0.077	**2.94E-35**	ARIC (European)		This study

https://doi.org/10.1371/journal.pgen.1008208.t002

would result in a premature stop codon (S1 Fig). We hypothesized that rs1939749 drives the *SLC22A24* expression levels and thus the minor allele results in lower transcript levels (S1 Fig). Interestingly, the minor alleles (A-alleles) of the sentinel SNPs (rs78176967 and rs151042642) are in complete linkage disequilibrium (D' = 1) with the major allele (G-allele) of rs1939749 and the minor allele (T-allele) of rs11231341 encoding Tyr501 (Fig 1F). The resultant haplotype, rs78176967:rs151042642:rs11231341:rs1939749 = A:A:T:G, which is present at ~5% haplotype frequency, is expected to retain function as there is no stop codon present at position 501 (T-allele) and no defective splicing (G-allele). The weaker p-value of the allele encoding Tyr501, rs11231341, is due to the fact that it may occur in two distinct haplotypes, one that is a loss of function due to the allele that is predicted to disrupt splicing (rs1939749), and one that retains function (no splicing defect) (Fig 1F). Overall, these results led to further studies designed to explore the functional role of SLC22A24 as a transporter of steroids and steroid conjugates.

Metabolomic studies and comparative structure modeling reveal SLC22A24 as an organic anion transporter

Metabolomic study. Large-scale metabolomic studies have been previously used in our laboratory and others to identify transporter substrates in an unbiased manner [31–33]. The technological advances of state-of-the-art metabolomic methods have allowed for increased sensitivity and accuracy in the quantification of small molecules, elevating the number of detectable endogenous metabolites achieving a Tier 1 standard for identification to over 700 [34] (S3 Table). In this study, we transiently transfected HEK293 Flp-In cells with either empty vector (EV) or *SLC22A24*. *SLC22A1* was transiently transfected as a positive control (S3 Table). We then incubated the cells with fetal bovine serum. Results in S3 Table show that many known metabolite substrates of SLC22A1 were significantly different between SLC22A1 and EV transfected cells, for example known substrates thiamine and acylcarnitines [31, 35]. A total of 50 metabolites reached the threshold of significance for difference in abundance between EV cells and SLC22A24 expressing cells (Fig 2A and 2B). Among the top 10 metabolites, four were bile acids (taurocholic acid, taurodeoxycholic acid, glycocholic acid and taurochenodeoxycholic acid), which have a steroid backbone, and two were dicarboxylic acids (3-hydroxy-3-methylglutarate and succinic acid) (Fig 2). Steroids and steroid conjugates were not above the threshold of detection in the cellular metabolomic analysis (S3 Table), potentially because of intrinsically low levels of endogenous steroids in the cell culture media containing fetal bovine serum [36]. Although most of the metabolites were present in greater abundance in the SLC22A24 expressing cells compared with EV cells, a few were present at lower abundance. Other anionic metabolites among the top 50 had monocarboxylic acid, phosphate, or sulfate moieties (S3 Table).

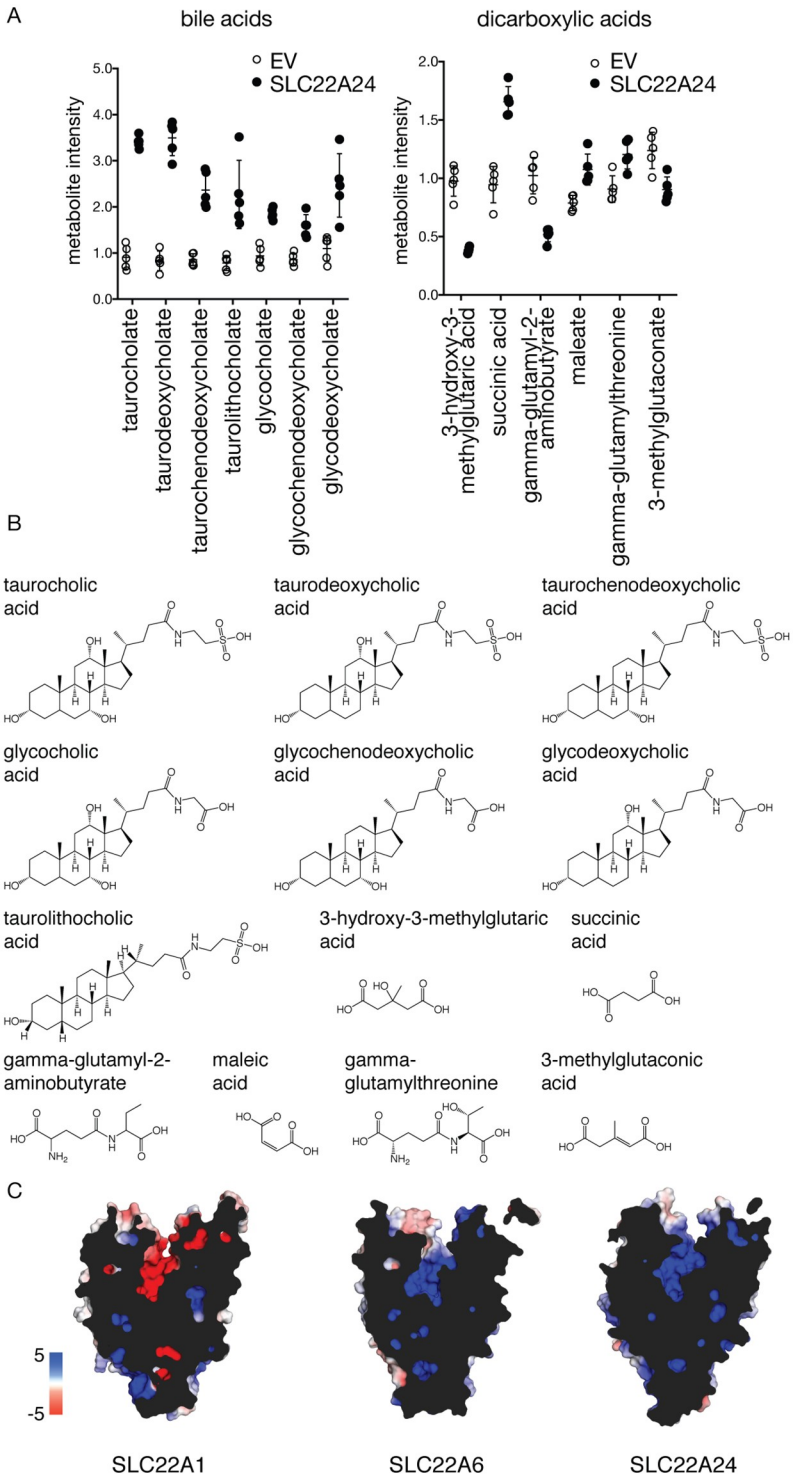


Fig 2. Metabolomic study and comparative structure models support the role of SLC22A24 as an anion transporter. (A) Cellular metabolite levels are significantly different between HEK293 Flp-In cells transiently transfected with empty vector (EV) or *SLC22A24* for seven bile acids and six dicarboxylic acids among the top 50 metabolites. Underlying data are provided in S1 Data. (B) Chemical structures of the seven bile acids and six dicarboxylic acids. (C) Surface representations of the calculated electrostatic potential (-5 to 5 kT e^{-1}) of the human SLC22A1, a characterized cationic transporter, the human SLC22A6, a characterized anionic transporter, and the human SLC22A24. The cross-section allows for visual inspection of the predicted binding pocket.

https://doi.org/10.1371/journal.pgen.1008208.g002

Comparative structure modeling. Due to the lack of available crystal structures for most SLC transporters, we modeled the human SLC22 members using the crystal structure of the human GLUT3 (SLC2A3) as a template. We hypothesized that despite the low sequence identity, the alignment would lead to an accurate assignment of residues along the helices, thus allowing an analysis of the electrostatic potential within the orthosteric site. As a control, we first modeled SLC22A1, a well-characterized organic cation transporter, and SLC22A6, an organic anion transporter (Fig 2C). A visual inspection of the electrostatic potential within the pocket was in agreement with the charges of known substrates. That is, the dominant charge within the SLC22A1 pocket was negative and within the SLC22A6 pocket was positive, consistent with binding and translocation of cationic and anionic substrates, respectively. The extension of the same methodology to SLC22A24 showed the pocket to carry an overall positive charge, implying an affinity for anionic compounds, and consistent with our metabolomic findings (Fig 2).

A diverse set of *in vitro* transporter assays supports the role of SLC22A24 as an organic anion transporter

We performed an uptake screen of 21 radiolabeled canonical substrates of various SLC22 family members, spanning cations, anions, and zwitterions. The experiments confirmed that SLC22A24 is a transporter of select organic anions (S4 Table).

Cells transiently transfected with *SLC22A24* accumulated steroid anionic compounds, including [^{3}H]-taurocholic acid (TA) (10–20 fold over empty vector control), [^{3}H]-glycocholic acid (GA) (10–20 fold), [^{3}H]-estrone sulfate (ES) (5–10 fold), [^{3}H]-estradiol-17β-glucuronide (EG) (3–5 fold), and [^{3}H]-androstanediol glucuronide (AG) (5–15 fold) (Fig 3A). The uptakes of [^{3}H]-estrone sulfate, [^{3}H]-estradiol-17β-D-glucuronide, [^{3}H]-androstanediol glucuronide, [^{3}H]-taurocholic acid and [^{3}H]-glycocholic acid were time dependent, reaching plateaus at approximately 5 min (S2 Fig). Kinetic parameters were evaluated at 1 min, except the uptake rate of taurocholic acid, which was evaluated at 2 min. The uptake kinetics of [^{3}H]-estrone sulfate, [^{3}H]-estradiol-17β-D-glucuronide, [^{3}H]-taurocholic acid, and [^{3}H]-glycocholic acid exhibited saturable characteristics at higher concentrations, with Km values (mean ± SD) of 8.6 ± 0.3 μM, 17.5 ± 0.2 μM, 10.5 ± 1.1 μM, and 33.4 ± 2.5 μM, respectively (Fig 3B, S5 Table). Similar values were observed for SLC22A8 (S5 Table). The kinetic parameters of androstanediol-3-glucuronide could not be determined accurately in SLC22A24 expressing cells due to a solubility limit of 300 μM (S3 Fig). Using the plateau at 300 μM as an approximation, the Km of androstanediol glucuronide for SLC22A24 was estimated to be 741 ± 253 μM, which is higher than for SLC22A8.

Next, we characterized the uptake of monocarboxylic and dicarboxylic acid derivatives, which are known substrates of other anion transporters, including several prescription drugs. Among the compounds that were identified as substrates of SLC22A24, two were dicarboxylic acids (i.e. succinic acid and 6-carboxyfluorescein, 6-CF) and two were monocarboxylic acids (i.e. 4',5'-dibromofluorescein, DBF, and 2'7'-dichlorofluorescein, DCF) (Fig 3C). However, other anion substrates of SLC22A6, SLC22A7, or SLC22A8, were determined not to be substrates of SLC22A24. Those include *para*-aminohippuric acid (PAH), uric acid, orotic acid, alpha-ketoglutaric acid and tetradecanedioic acid (S4 Table). The prescription drug, methotrexate, and the mycotoxin, ochratoxin A, showed significant uptake by SLC22A24 expressing over EV cells (Fig 3D), whereas tenofovir and adefovir, which have phosphate moieties, were not substrates of the transporter (S4 Table). Further, none of the tested cationic and zwitterionic substrates of SLC22A1-SLC22A5 showed significant (>1.5 fold) accumulation in SLC22A24 expressing cells compared to EV cells (S4 Table).

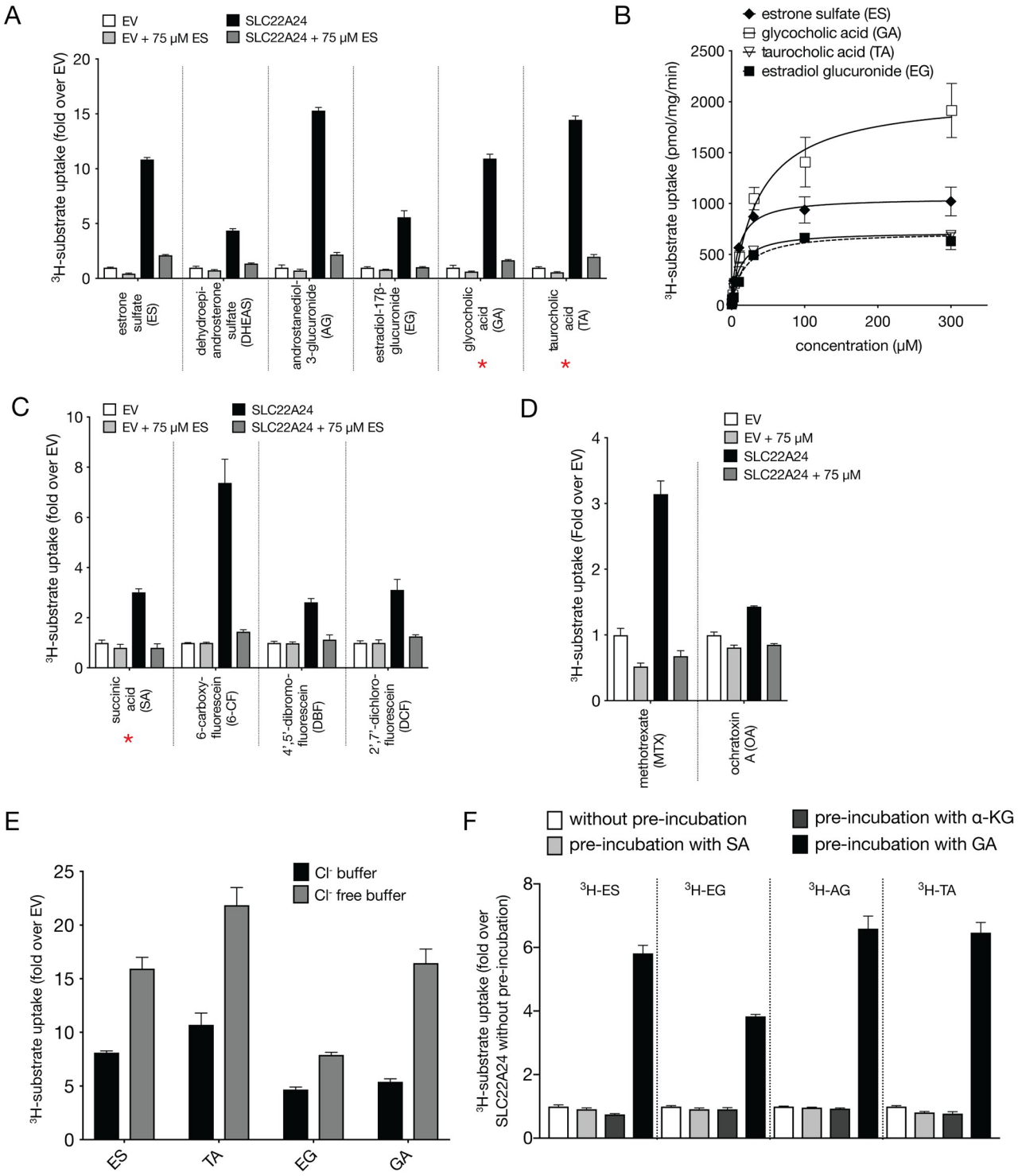


Fig 3. Uptake and inhibition studies of various anions in HEK293 Flp-In cells stably expressing SLC22A24. (A) Uptake of steroid sulfate or glucuronide (ES, DHEAS, AG, EG) conjugates and bile acids (GA, TA). Uptake of these anions was significantly different between EV and *SLC22A24* transfected cells. Uptake was significantly reduced in the presence of excess unlabeled estrone sulfate. Figure shows a representative plot from one experiment (mean ± S.D. from three replicate wells). The experiments were repeated at least one time and showed similar results. (B) Kinetics of uptake of four anions for SLC22A24. The uptake rate was evaluated at 1 minute, except for taurocholic acid, which was evaluated at 2 min. Each point represents the uptake (mean ± S.D.) in cells stably expressing SLC22A24, normalized to the empty vector control. The data were fit to a Michaelis–Menten equation. The figure shows a representative plot from one experiment. The mean and S.D. of the kinetic parameters from two experiments are shown in S5 Table. The time-dependent uptakes of selected anions are shown in S2 Fig. (C, D) Uptake of anions with carboxylic acid and dicarboxylic

acid groups. Uptake of these anions was significantly different between cells stably expressing SLC22A24 and the empty vector control cells. Uptake was reduced significantly in the presence of excess unlabeled estrone sulfate. Figure shows a representative plot from one experiment (mean ± S.D. from three replicate wells). The experiments were repeated at least one time and showed similar results. (E) Uptake of four anions in HEK293 Flp-In cells stably expressing SLC22A24 with chloride-containing and chloride-free buffer. Each bar represents uptake (mean ± S.D. from three replicate wells) in the SLC22A24-transfected cells normalized to the empty vector control. The experiments were repeated at least one time and showed similar results. (F) Trans-stimulation of SLC22A24–mediated ES, EG, AG and TA uptake. The uptake of ES, EG, AG and TA were trans-stimulated by separately preloading the cells with 2 mM of succinic acid, alpha-ketoglutrate acid, or glutaric acid for 2 h. After preloading, the uptake of the anions was measured after 10 min. Data are means ± S.D. The experiments were repeated one time, in triplicate, and show similar results. All experiments were normalized by setting the uptake of SLC22A24-expressing cells without preloading to 100%. Underlying data are provided in S1 Data.

https://doi.org/10.1371/journal.pgen.1008208.g003

Collectively, the data indicate that SLC22A24 transports various organic anions, including steroid compounds with anionic moieties (sulfates, carboxylic acids, glucuronides) and non-steroid compounds with dicarboxylic acid moieties.

SLC22A24 transport is chloride sensitive and transtimulated by glutaric acid

We examined the underlying transport mechanism of SLC22A24 in order to determine whether it relies on facilitative diffusion or on secondary active transport of ions. Uptake of [^{3}H]-estrone sulfate was not affected by pH between 5.5 to 8.5 and was not dependent on the presence of Na^+ ions (S4 Fig). However, uptake of [^{3}H]-estrone sulfate, as well as that of other steroid conjugates, was significantly affected by the presence of Cl^- ions. Cl^-free buffer nearly doubled the uptake of all tested compounds (Fig 3E). In addition, we tested trans-stimulation of SLC22A24-mediated uptake by preloading the cells with succinic acid, alpha-ketoglutaric acid, and glutaric acid, which are known to trans-stimulate other organic anion transporters [37–39]. Among the three, only glutaric acid significantly increased the uptake of [^{3}H]-estrone sulfate, [^{3}H]-estradiol glucuronide, [^{3}H]-androstanediol glucuronide, and [^{3}H]-taurocholic acid (Fig 3F).

SLC22A24 is inhibited by anionic compounds from different chemical classes

The primary goal of the inhibition studies was to confirm that solutes identified in the metabolomic studies were ligands of SLC22A24 and to identify additional ligands of the transporter. The secondary goal was to determine whether known inhibitors of other organic anion transporters in the SLC22 family were inhibitors of SLC22A24. A total of 85 compounds were selected for inhibition studies of SLC22A24-mediated [^{3}H]-taurocholic acid uptake (S6 Table, Fig 4A). These included: (i) 4 steroid metabolites identified in GWAS at $p<1x10^{-6}$; (ii) 20 known inhibitors or substrates of other organic anion transporters; (iii) 17 steroids or bile acids which are used clinically; (iv) 8 dicarboxylic acid metabolites identified in our metabolomic study; and (v) 20 steroid conjugates.

At 100 μM, 43 of the 86 test-compounds inhibited the SLC22A24-mediated [^{3}H]-taurocholic acid uptake by >50%. All 16 bile acids, 5 of the 15 unconjugated steroids, 15 of the 20 steroid conjugates, and 12 of the 22 other anionic non-steroidal compounds inhibited taurocholic acid uptake by >50% (Fig 4A, S6 Table). Many sulfate conjugates including steroid sulfates inhibited taurocholic acid uptake. Some non-steroidal compounds known to inhibit organic anion transporters also inhibited SLC22A24-mediated taurocholic acid uptake potently (e.g. losartan, rosiglitazone, lesinurad, benzbromarone) or partially (e.g. probenecid, bumetanide, ketoprofen). Three steroid metabolites identified in the GWAS, progesterone, androsterone glucuronide, and etiocholanolone glucuronide inhibited taurocholic acid uptake >50%. Consistent with these results, SNPs in *SLC22A24* were associated with progesterone,

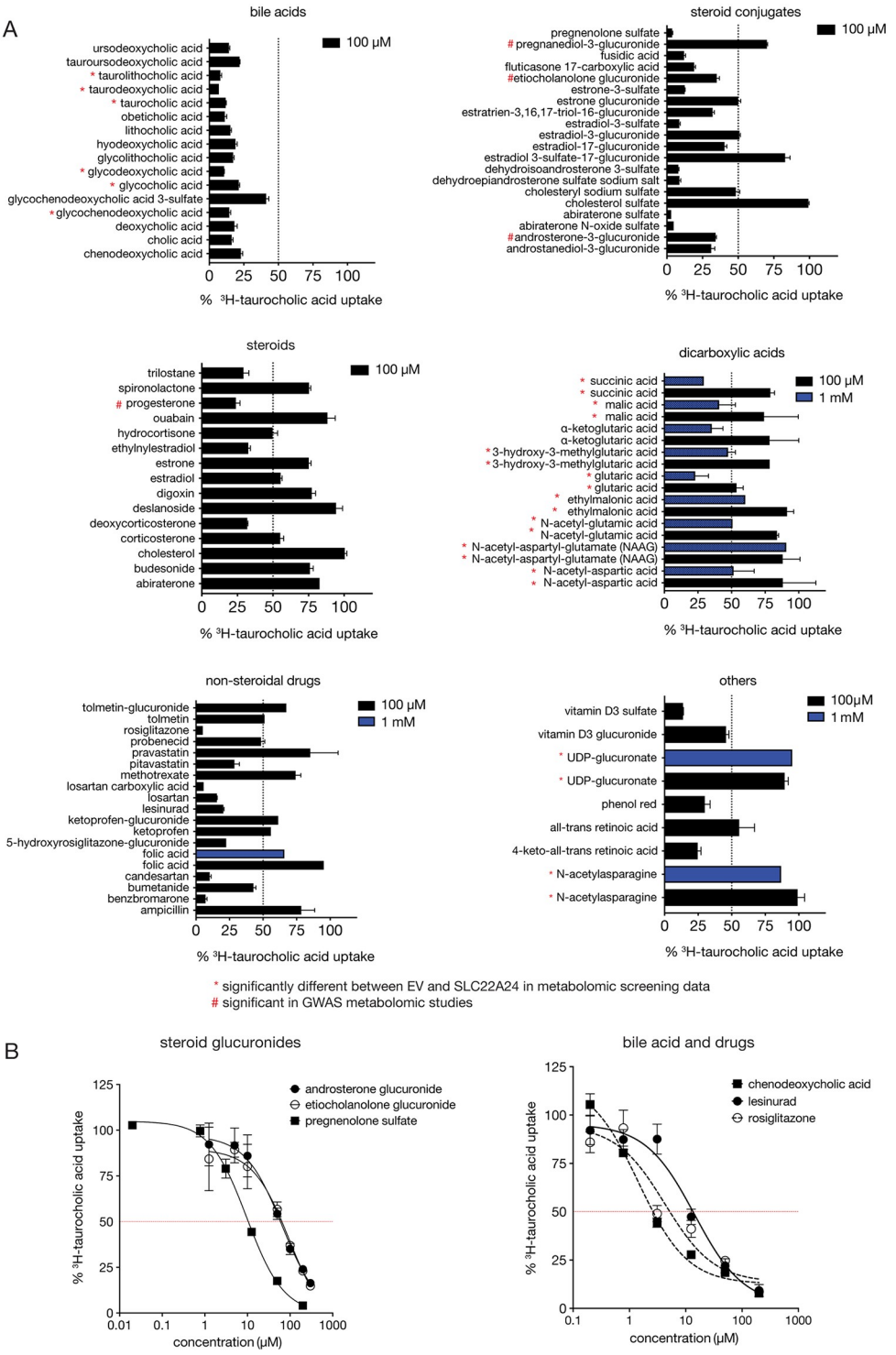


Fig 4. Inhibition studies of various anions in HEK293 Flp-In cells stably expressing SLC22A24. (A) Inhibition of [^{3}H]-taurocholic acid uptake (with 1 µM unlabeled taurocholic acid) in cells stably expressing SLC22A24 by various bile acids, steroids, steroid conjugates, dicarboxylic acids, non-steroidal drugs, and other substances. All compounds were screened at 100 µM, and for some compounds, also at 1 mM. Uptake of [^{3}H]-taurocholic acid was stopped after 10 min. Values represent the mean ± S.E.M. (from at least one experiment with three replicate wells each time). *Compounds that are statistically significant in the metabolomic analysis. #Steroidal compounds that are statistically significant in the genome-wide association studies. (B) Inhibition plots for SLC22A24-mediated taurocholic acid uptake in HEK293 Flp-In cells stably expressing SLC22A24. Cells were incubated with HBSS buffer containing taurocholic acid (1 µM) for 10 min with or without various concentrations of the compounds. Values are presented as

mean ± S.D. of taurocholic acid uptake from three replicate wells determined in a single experiment. The experiments were repeated once and similar results were obtained. Representative curves of the SLC22A24-mediated taurocholic acid uptake inhibition by steroid conjugates (progesterone sulfate, etiocholanolone glucuronide), bile acid, and drugs (chenodeoxycholic acid, lesinurad, rosiglitazone). Underlying data are provided in S1 Data. See Table 3 for the IC_{50} values.

https://doi.org/10.1371/journal.pgen.1008208.g004

androsterone glucuronide and etiocholanolone glucuronide at genome-wide significance levels ($p<5x10^{-8}$), whereas associations were weaker for pregnanediol-3-glucuronide ($p = 2x10^{-7}$), which was a weaker inhibitor of SLC22A24-mediated taurocholic acid uptake (Fig 4A) [8]. At a high concentration (1 mM), five of the nine dicarboxylic acids inhibited taurocholic acid uptake by approximately 50% (Fig 4A). However, UDP-glucuronate and N-acetylasparagine did not inhibit taurocholic acid uptake even at 1 mM.

Ten compounds were selected for further IC_{50} studies (Table 3). We specifically focused on the steroids and steroid conjugates reported in GWAS and the steroid conjugates with the strongest inhibition in the initial screen. Two bile acids, chenodeoxycholic acid and ursodeoxycholic acid, were selected because they are used clinically to treat gallstone and related cholestatic liver disease [40, 41]. Two non-steroidal drugs, rosiglitazone and lesinurad, were selected because of their strong inhibition in the initial screen. Pregnenolone sulfate, chenodeoxycholic acid, and rosiglitazone exhibited potent IC_{50} values ($IC_{50} < 5$ μM) (Fig 4B). Progesterone, ursodeoxycholic acid, and lesinurad were also inhibitors with $IC_{50} < 20$ μM. Finally, the three steroid glucuronide conjugates, androsterone glucuronide, etiocholanolone glucuronide, and androstanediol-3-glucuronide, achieved IC_{50} values for inhibition of SLC22A24-mediated taurocholic acid uptake of 45 ± 29 μM, 57 ± 36 μM and 21 ± 11 μM, respectively (Table 3). The IC_{50} values of androsterone glucuronide and etiocholanolone glucuronide are relatively similar to the reported concentrations of these metabolites in the urine (ranges from 1–33 μM for androsterone glucuronide and 0.9–20.4 μM for etiocholanolone glucuronide) [42, 43].

SLC22A24 orthologs exhibit substrate selectivity differences

Phylogenetic analysis. First, we determined the relationships between the human SLC22A24 with other human, mouse, and rat SLC22 family members with known substrates (Fig 5A). In the phylogenetic analysis we also included genes within the locus between

Table 3. Inhibition potencies of steroid, steroid conjugates, bile acids and drugs for SLC22A24-mediated taurocholic acid uptake. The experiments were repeated at least once to obtain IC_{50} mean ± SD (the inhibition concentration at 50% taurocholic acid uptake).

Compound	Class	IC_{50} [μM ± SD]
Progesterone	Steroid	7.4 ± 3.0
Pregnenolone sulfate*	Steroid conjugate	1.4 ± 0.1
5α-Androstan-3α-ol-17-one glucosiduronate (androsterone glucuronide)		45 ± 29
5β-Androstan-3α-ol-17-one glucosiduronate (etiocholanolone glucuronide)*		57 ± 36
5α-Androstan-3α,17β-diol 3-glucuroside (androstanediol glucuronide)		21 ± 11
Pregnanediol-3-glucuronide		> 200
Chenodeoxycholic acid*	Bile acid	2.6 ± 1.0
Ursodeoxycholic acid		7.6 ± 1.2
Lesinurad*	Drug	13 ± 6.1
Rosiglitazone*		3.3 ± 1.4

*Representative IC_{50} curves for these compounds are shown in Fig 4B.

https://doi.org/10.1371/journal.pgen.1008208.t003

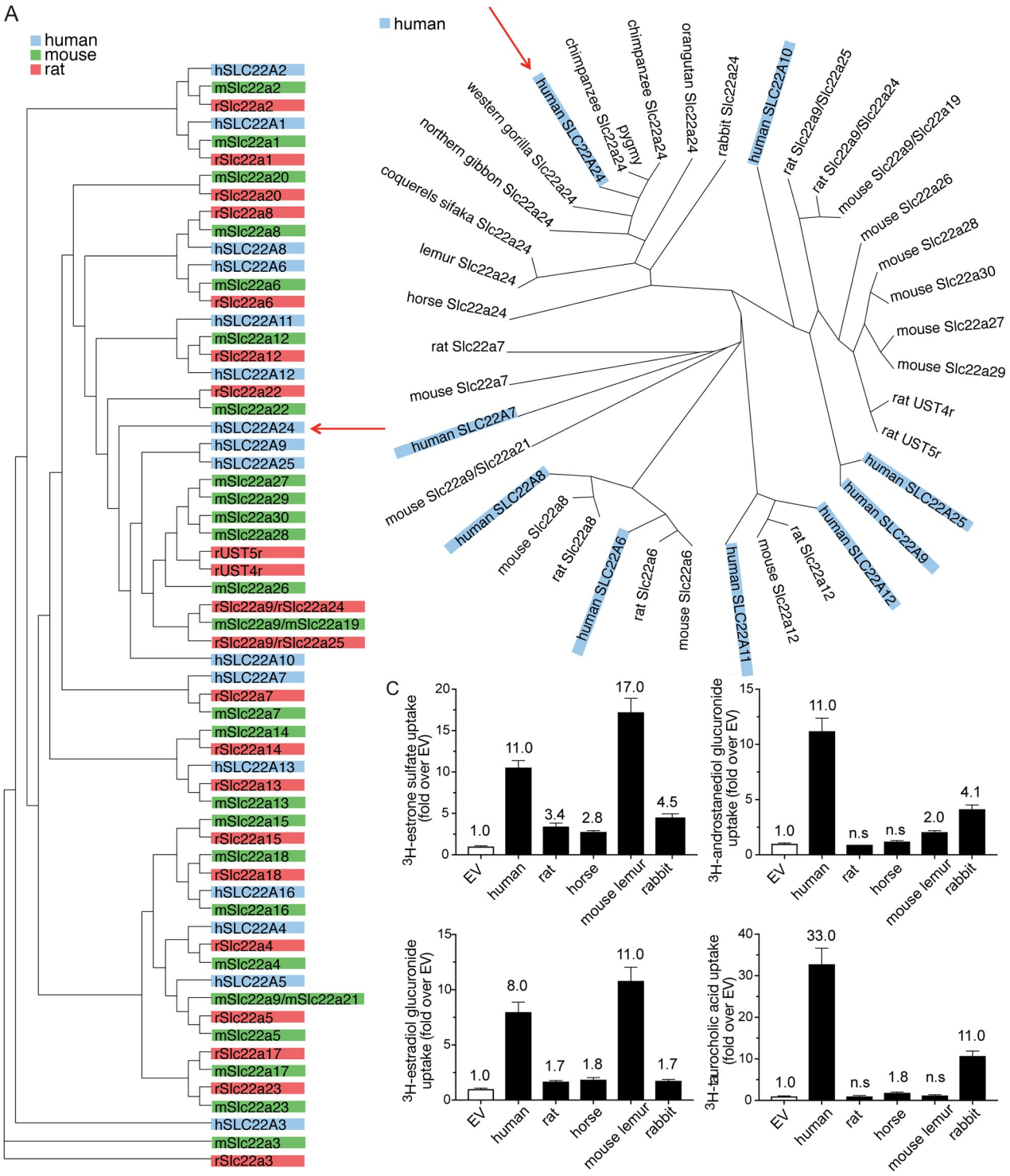


Fig 5. Phylogenetic tree analyses and uptake of anions by species orthologs of SLC22A24. (A) Sequence comparison of human, mouse, and rat protein sequences for all characterized SLC22 family members. Proteins are colored by organism. Red arrow points to the human SLC22A24. (B) Sequence comparison of characterized human anion transporters and select mammalian orthologs. Red arrow points to the human SLC22A24. (C) Fold uptake of four substrates by species orthologs of SLC22A24 and rat Slc22a9/Slc22a24 over empty vector control. Uptake of estrone sulfate, androstanediol glucuronide, estradiol glucuronide, and taurocholic acid was performed in HEK293 Flp-In cells transiently expressing *SLC22A24* from human, mouse lemur, rabbit and horse or *Slc22a9/a24* from rat. Uptakes were performed for 10 minutes. The number above the bar is fold uptake over EV cells from one experiment (mean ± S.D., from four replicate wells). The experiments were repeated one time and similar results were obtained. Underlying data are provided in S1 Data.

https://doi.org/10.1371/journal.pgen.1008208.g005

SLC22A8 and SLC22A9, known as the "unknown substrate transporters (USTs)" [44]. Notably, the number of genes annotated as USTs in mouse and rat is different from human. Mouse USTs include Slc22a26, a27, a28, a29, a30, a9/a19. Rat USTs include Slc22a9/a24, a9/a25, Ust4r, Ust5r. This analysis showed that human SLC22A24 lacks close human paralogs and rodent orthologs, unlike some well-characterized SLC22 family members such as SLC22A1 and SLC22A2 (Fig 5A).

Data from Ensembl [45, 46] revealed that SLC22A24 1:1 orthologs were present in some, but not all mammals, and were absent in non-mammalian vertebrates (e.g., fish, reptiles and birds). Fig 5B shows the phylogenetic analysis of ten SLC22A24 1:1 orthologs in mammals along with other relevant organic anion transporters in the SLC22 family (see also S5 Fig). Included in the figure are SLC22A24 orthologs from apes (orangutan, chimpanzee, gorilla and gibbon), lemur (mouse lemur and coquerels sifaka), rabbit, and horse. We compared the function of the human SLC22A24 with ortholog representatives from the different branches of the phylogenic tree (Fig 5C), that of the horse, mouse lemur, and rabbit. We also included the rat Slc22a9/Slc22a24, previously known as Slc22a19 and Slc22a9, but most recently assigned the gene name *Slc22a24* [47]. All five species (human, horse, mouse lemur, rabbit, and rat) showed significant uptake of estrone sulfate (>2.5 fold) compared to EV transfected cells (Fig 5C). There were substrate specificity differences between the human SLC22A24 and the orthologs of the other species investigated with the largest differences between the rat orthologs and the other species. Notably, horse SLC22A24 significantly transports taurocholic acid (~2 fold), estrone sulfate (~3 fold), and estradiol-17-glucuronide (~2 fold), but not androstanediol glucuronide (~1.2 fold). Mouse lemur SLC22A24 transports estrone sulfate (~17 fold), androstanediol glucuronide (~2 fold) and estradiol-17-glucuronide (~11 fold) but not taurocholic acid. Rabbit SLC22A24 significantly transports taurocholic acid (~11 fold), estrone sulfate (~5 fold), androstanediol glucuronide (~4 fold), and estradiol-17-glucuronide (~2 fold). In contrast, rat Slc22a9/Slc22a24, which has the lowest sequence similarity to human protein of the four species, does not significantly transport taurocholic acid or androstanediol glucuronide.

We also determined the plasma membrane expression of each of the above orthologs in HEK293 cells using a GFP-tag to ensure the proper localization of the expressed proteins (S6 Fig). All five orthologs were expressed on the plasma membrane, but the expression levels were variable. More pronounced plasma membrane expression was observed for the rabbit and human orthologs. The horse, rat, and the mouse lemur orthologs, were also present on the plasma membrane, but had a more diffuse intracellular expression (S6 Fig).

Two SLC22A24 isoforms exhibit no significant differences in their substrate selectivity, while SLC22A24 p.Tyr501Ter variant exhibits no protein expression or transport function

According to UniProt, Ensembl, and NCBI RefSeq, the SLC22A24 transcript is identified in three major isoforms. Specifically, ENST00000612278.4 (552 amino acids, NM_001136506, A0A087WWM3), identified in this paper as the reference SLC22A24, ENST00000417740.5 (551 amino acids, no known NCBI reference, C9JC66), identified as SLC22A24-551, and ENST00000326192.5 (322 amino acids, NM_173586, Q8N4F4), identified as SLC22A24-322. The short SLC22A24-322 isoform was omitted from further analysis due to its lack of transport activity, expression from GFP fusion construct, and absence of key structural motifs (see S7 Fig). Transcripts of the two isoforms (SLC22A24 and SLC22A24-551) were detected using PCR and the sequences were confirmed. The SLC22A24-551 isoform exists due to the presence of two potential AG acceptor sites at the splice junction between exons 9 and 10, (indicated as position "1" or "2" on S8A Fig), thus resulting in two different splice variants. The splice

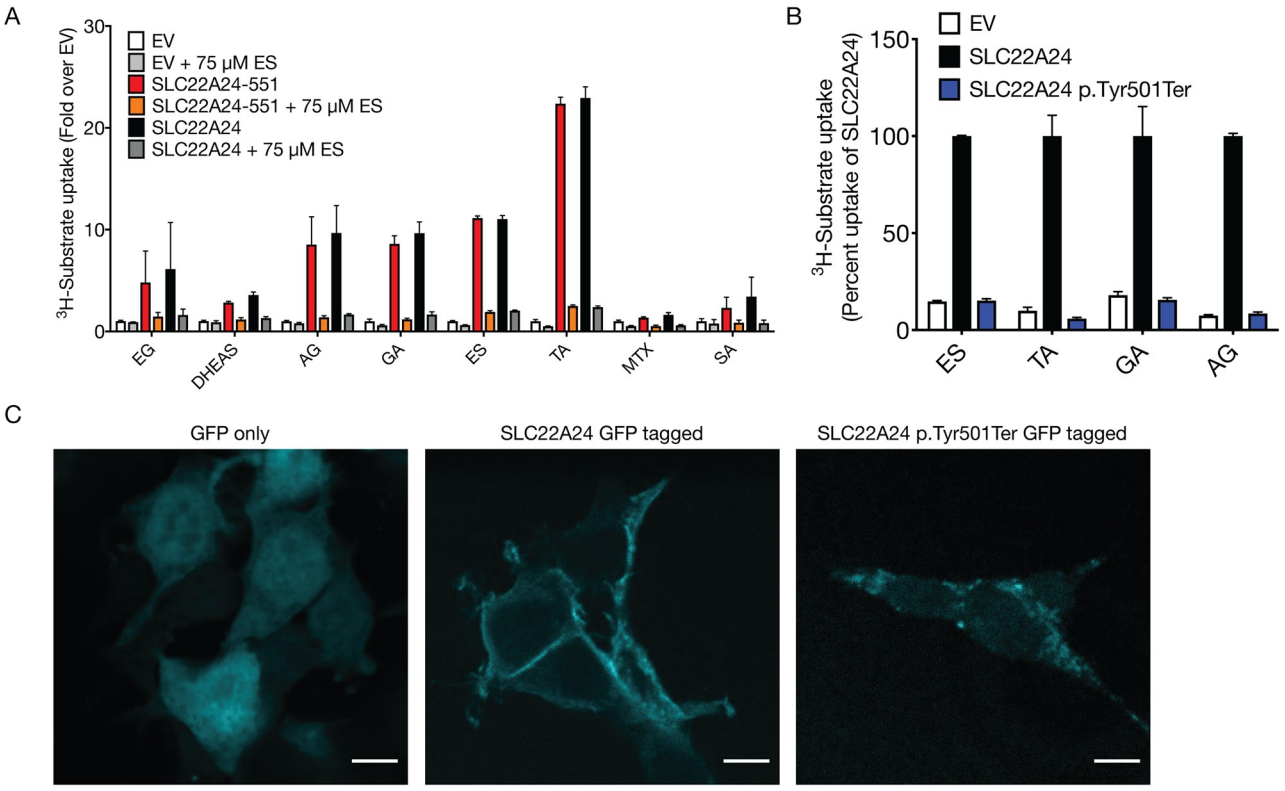


Fig 6. Uptake of various anions by two SLC22A24 isoforms and the effect of a common nonsense variant, p.Tyr501Ter on uptake and plasma membrane localization. (A) Fold uptake over empty vector for different substrates. HEK293 Flp-In cells were transiently transfected with different constructs containing *SLC22A24* or *SLC22A24-551*. The experiments were repeated at least one time. Data represents means ± S.D. of two independent experiments. Underlying data are provided in S1 Data. (B) Transport activity for different substrates was compared in HEK293 Flp-In cells transiently transfected with *SLC22A24* or *SLC22A24* p.Tyr501Ter. The uptake of SLC22A24-expressing cells was set to 100%. Underlying data are provided in S1 Data. Data represents means ± S.D. from three wells. The experiments were repeated one time and similar results were obtained. (C) Membrane localization of SLC22A24 and SLC22A24-Y501Ter. HEK293 cells were transiently transfected with GFP fusion constructs and visualized by confocal microscopy.

https://doi.org/10.1371/journal.pgen.1008208.g006

variants differ in two amino acids in position 533 and 534 (S8B Fig). Functional characterization of SLC22A24 and SLC22A24-551 indicated no significant differences in their uptake function. In particular, a similar fold-uptake over control between the two isoforms was observed for all representative steroid glucuronides, steroid sulfates, bile acids, methotrexate and succinic acid (Fig 6A).

Next, we characterized the nonsense variant in SLC22A24, which is associated at genome-wide significance levels with androsterone glucuronide and etiocholanolone glucuronide (Fig 1). A survey of the 1000 Genomes Project Phase 3, revealed that the G-allele of rs11231341, coding for p.Tyr501Ter, is the major allele, ranging from 51% in Mende in Sierra Leone to 92% in Peruvian in Lima, Peru [48]. In addition, the G-allele is the major allele (80%) in individuals from the Greater Middle East (GME) (The GME Variome Project, http://igm.ucsd.edu/gme/). Cells transiently transfected with SLC22A24 p.Tyr501Ter exhibited no transport function (Fig 6B). Consistent with the uptake studies, GFP-tagged SLC22A24 localized to the membrane, while GFP-tagged SLC22A24 p.Tyr501Ter did not (Fig 6C). Further, plasma membrane protein expression was evaluated using Western blots for HEK293 Flp-In cells stably expressing SLC22A24 or transiently transfected with either empty vector only (EV, pCMV6-Entry), SLC22A24, or SLC22A24 p.Tyr501Ter (S9 Fig). All samples expressed the

positive control, Na^+/K^+-ATPase. SLC22A24 was expressed in both the stable cell sample as well as the transiently transfected cells. SLC22A24 p.Tyr501Ter was not detected on the plasma membrane of the transiently transfected cells, despite having detectable mRNA levels (S9 Fig).

Transcriptomic and proteomic studies reveal low and variable expression levels of multiple isoforms of SLC22A24

Transcriptomic studies. PCR identified the transcripts to be expressed at variable levels in the 14 tissues tested (pooled from various samples, see S7 Table), in addition to kidney and liver (Fig 7A). SLC22A24 and SLC22A24-551 were found at moderate levels in nine tissues including total brain, liver, skeletal muscle, kidney, pancreas, colon, spleen, testis, and thymus (Fig 7A). Interestingly, the Kidney Interactive Transcriptomics database, which is a repository for single cell RNAseq data from human kidney, shows SLC22A24 to be expressed specifically in the proximal tubule segment 3 (S3), and not in other regions of the human kidney (http://humphreyslab.com/SingleCell/search.php, Fig 7B)[49], providing support for low overall transcriptomic expression in the kidney.

Proteomic studies. Using LC-MS/MS Selective Reaction Monitoring (SRM) approach, we evaluated protein levels of the two SLC22A24 isoforms in 20 renal cortical samples from self-reported African Americans (age [mean ± SD]: 13 ± 8 years; 50% males). SRM is a sensitive mass spectrometry technique that is capable of accurate protein detection and quantification [50]. Three peptides were designed to detect and quantify the protein abundance (S10 Fig). Peptide 1 detected both SLC22A24 and SLC22A24-551, while peptides 2 and 3 were specific to SLC22A24 and SLC22A24-551, respectively. Surprisingly, isoforms SLC22A24 and SLC22A24-551 exhibited a mutually exclusive expression pattern (S8 Table), that is, each individual sample expressed only one of the two isoforms. Genotyping 20 renal cortical samples from African Americans, we identified the c.1503-G allele to have an allele frequency of 67.5% (Hardy-Weinberg p-value = 0.4), consistent with known allele frequencies in publicly available sources. Ten c.1503-GG, seven c.1503-GT, and three c.1503-TT were used for the SRM analysis (*SLC22A24* c.1503-GG codes for SLC22A24 p.Tyr501Ter homozygotes). In line with the stop codon at amino acid position 501, the SRM analysis showed that the renal cortical tissues with the genotype c.1503-GG have no detectable protein abundance using peptides 2 and 3 (S8 Table).

Discussion

Steroids including sex steroids and glucocorticosteroids have pleiotropic actions in reproductive and other tissues in the body and impact numerous biological processes [1, 53, 54]. In particular, they influence body composition, glucose and lipid metabolism, cardiovascular and immunological functions, and reproductive fitness [54]. Accordingly, there is a vast literature focused on the study of the action, regulation, metabolism and disposition of steroids [1]. GWAS have revealed a number of genes that are involved in steroid metabolism and hormone binding as the major determinants of interindividual variation in the levels of many steroids and steroid conjugates [2–7, 55]. Intrigued by the association of a novel locus and an orphan transporter, SLC22A24, with metabolite levels in metabolomic GWAS, we performed experimental studies along with a computational analysis to functionally characterize the transporter and further understand its physiological role. Our studies led to four major findings. First, SLC22A24 is an organic anion transporter, consistent with multiple GWAS focused on metabolites, and validated through comparative structure models, cell-based metabolomic screens, and functional studies. Second, SLC22A24 is conserved among higher order primates, but has diverged within lower level primates and other mammals. Third, two functional isoforms of

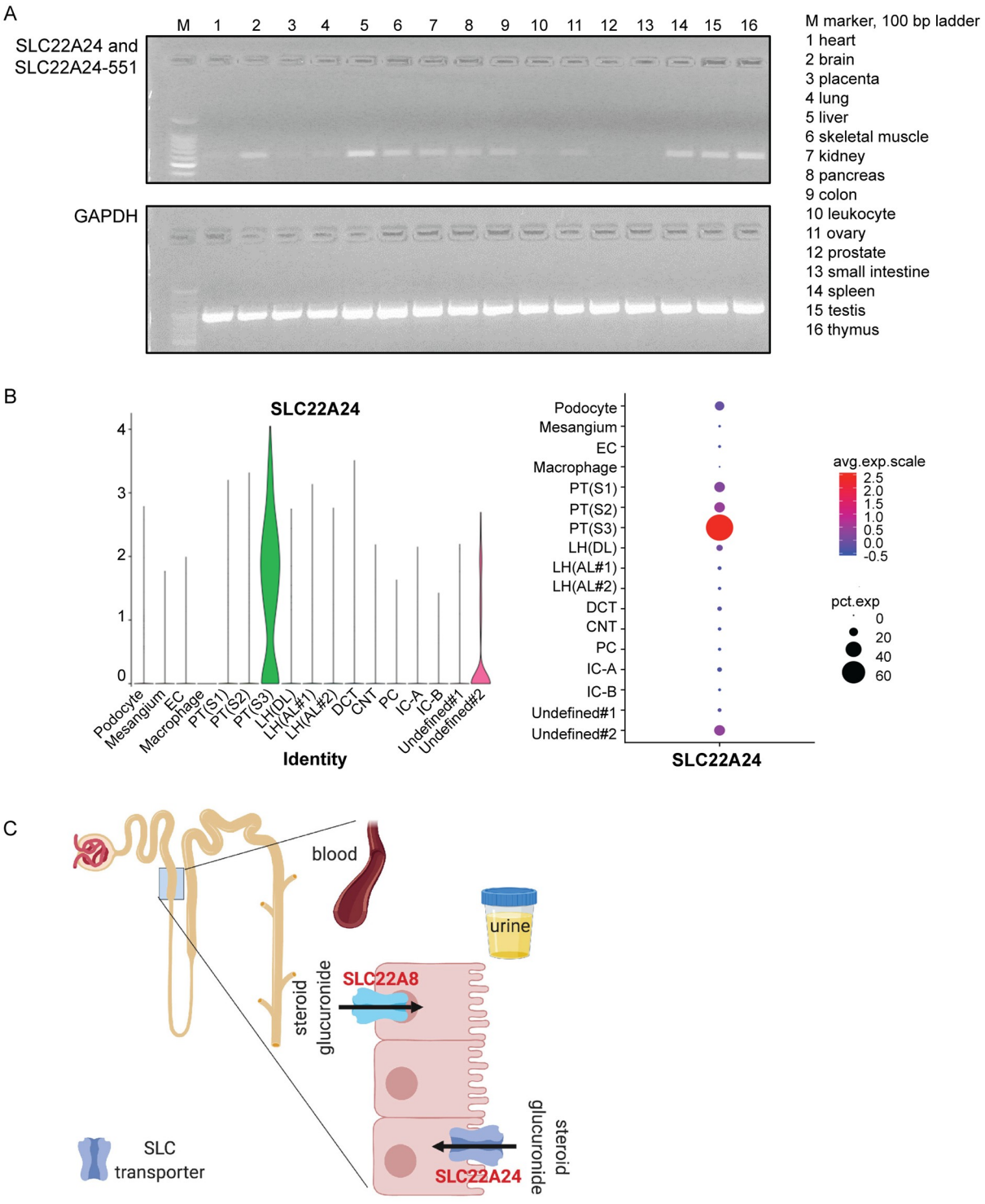


Fig 7. Expression of *SLC22A24* mRNA transcripts in human tissues and cells, and a proposed model for its functional role in the disposition of steroid conjugates in the kidney. (A) Expression of *SLC22A24* mRNA in various pooled samples was detected by RT-PCR. 40 and 30 cycles were used for amplifying *SLC22A24* and *GAPDH* mRNA, respectively. (B) *SLC22A24* expression from human kidney single cell datasets (Healthy Human Adult Kidney (Complete: 4,524 nuclei)) were obtained from Kidney Interactive Transcriptomics (http://humphreyslab.com/SingleCell/)[51]. Two different plots (i and ii) from the online analyzer for kidney single cell datasets, show the regions of the kidney and the abundance of *SLC22A24* mRNA transcripts. EC, endothelial cell; PT, proximal tubule; LH(AL), loop of Henle ascending loop; LH(DL), loop of Henle descending loop; DCT, distal convoluted tubule; CNT, connecting tubule; PC, collecting duct-principal cell; IC-A, intercalated cell type A; IC-B, intercalated cell type B. Overall, the plots indicate that *SLC22A24* is selectively expressed in the proximal

tubule segment 3 (PT(S3)). (C) Proposed mechanism of steroid glucuronide handling in the kidney by SLC22 transporters. The scheme illustrates the proposed hypothesis that SLC22A24 is on the apical membrane of the proximal tubule and functions in the reabsorption of steroid glucuronides. Conversely, SLC22A8, localized to the basolateral membrane, functions in the secretion of steroid glucuronides[52]. Future studies are needed to precisely localize the transporter to the apical or basolateral membrane. This scheme was created with BioRender (https://biorender.com/).

https://doi.org/10.1371/journal.pgen.1008208.g007

SLC22A24 are detectable in the human kidney, and the transporter has a highly specific expression pattern in the nephron, localized to segment 3 of the proximal tubule. Fourth, the transporter has a common nonsense variant, p.Tyr501Ter (rs11231341), which is found at high allele frequencies in human populations, and is associated with low steroid and steroid conjugate levels ($p<5x10^{-8}$). Below we discuss each of the major findings in detail.

The first major finding of the study is that SLC22A24 is an organic anion transporter (Figs 2 and 3). Based on results the GWAS (Fig 1), metabolomic profiling and comparative structure modeling (Fig 2), we hypothesized and validated in our *in vitro* and cellular metabolomic studies that SLC22A24 preferentially transports organic anions. However, it does not transport a number of organic anions transported by other organic anion transporters in the SLC22 family (SLC22A6, SLC22A7 and SLC22A8). For example, we did not detect transport of uric acid, tenofovir, cGMP, or alpha-ketoglutaric acid, which are canonical substrates of SLC22 organic anion transporters (S5 Table) [20, 56]. Further, our results showed that SLC22A24 is Cl^-dependent (Fig 3E). A similar chloride dependency has been observed for URAT1 (SLC22A12), a major uric acid reabsorptive transporter in the kidney [57], and OAT4 (SLC22A11), which uses the downhill flux of chloride ions to move its organic anion substrates across the apical membrane of the renal proximal tubule [37, 58]. Similar to OAT1 (SLC22A6), OAT3 (SLC22A8) and OAT4 (SLC22A11), SLC22A24 was also trans-stimulated by glutaric acid [37, 38, 59, 60], but not significantly by succinic acid and alpha-ketoglutaric acid (Fig 3F). In contrast, OAT1 and OAT3 are trans-stimulated by alpha-ketoglutaric acid as well as glutarate, consistent with a broader selectivity of these two transporters in comparison to SLC22A24.

Our second finding is that many species, including lower level primates and rodents, do not have orthologs of SLC22A24. Phylogenetic analysis reveals the presence of SLC22A24 orthologs in apes, lemur, and certain other mammalian species, but none in rodents. Human *SLC22A24*, located on chromosome 11q12, falls within a cluster of other SLC22 family genes called the Unknown Substrate Transporters (USTs), consisting of *SLC22A9*, *SLC22A10*, and *SLC22A25*. Sequence comparison indicates that human SLC22A24 is distant from those in mice and rats (mouse USTs include Slc22a26, a27, a28, a29, a30, a9/a19; rat USTs include Slc22a9/a24, a9/a25, Ust4r, Ust5r) (Fig 5A, S5 Fig) and only 11 species have direct orthologs of the human SLC22A24 (see ref. [46], Ensembl database (version 95)). Previous studies have characterized the function of human liver specific SLC22A9 (OAT7)[25], several rodent USTs, including the rat kidney Slc22a9/a24 (Oat5) [61], rat kidney Slc22a9/a25 (Ustr1/Oat8) [62, 63], mouse kidney Slc22a9/a19 (Oat5) [64–66], and mouse liver and kidney Slc22a30 (Oat9) [67]. All of the above transporters transport steroid sulfate conjugates and some other anions, such as ochratoxin A. However, they do not transport and are not inhibited by steroid glucuronide conjugates, which are substrates and inhibitors of the human SLC22A24 [61, 63]. Consistent with these results, steroid glucuronides are not detectable in rat urine but are detectable in human urine [68]. Further, glucuronidated steroids are not detected in serum samples from rodents [69].

Our third finding is that the functional isoforms of the human SLC22A24, are expressed at lower levels in kidney and liver tissues compared to other known kidney-specific anion transporters from the SLC22 family, e.g., SLC22A6, SLC22A8, SLC22A11, and SLC22A12, as noted from the RNA-seq data generated by the Human Protein Atlas (HPA)[70] and the Genotype-

Tissue Expression (GTEx) consortium (see S10 Table). Though expressed at lower levels, the primary finding that genetic variants in SLC22A24 associate with plasma levels of several metabolites (S1 Table), suggests that the transporter plays a critical and specific role in determining systemic levels of metabolites. The narrower substrate selectivity of SLC22A24 as compared to other SLC22 organic anion transporters, SLC22A6 and SLC22A8, which are broadly selective, further supports a specific role for the transporter in the kidney. The Kidney Interactive Transcriptomics database (http://humphreyslab.com/SingleCell/search.php) shows SLC22A24 to be expressed specifically in S3 (Fig 7B) and not in the other regions of the human kidney, which suggests that SLC22A24 plays a specific role in the kidney in comparison with other renal organic anion transporters in the SLC22 family. The selective localization of SLC22A24 in the human kidney is consistent with the renal localization of rodent USTs, which are expressed in particular sub-regions of the kidney [61, 63, 65]. For example, rat Slc22a9 is expressed on the apical membrane of the collecting duct [63], and mouse Slc22a9/a19 is expressed on the apical membrane of the proximal tubules of the outer medulla [61, 65]. Collectively, the data suggest a specific role for SLC22A24 in the kidney, reminiscent of USTs in rodents, which do not express SLC22A24.

Our fourth finding is that the transporter has a common nonsense variant, p.Tyr501Ter (rs11231341), which is found at high allele frequencies in human populations, and is a major determinant of inter-individual variation in plasma levels of certain steroid and steroid conjugates ($p<5x10^{-8}$). Premature stop codons sometimes trigger nonsense-mediated decay (NMD), which in this case might cause a rapid degradation of the SLC22A24 transcripts and result in the observed lack of detectable protein signal in the kidney samples from individuals who are homozygous for p.Tyr501Ter (S8 Table). The overall low mRNA transcript levels of SLC22A24 in publicly available databases (which either average expression levels from multiple samples or report results of pooled samples) may be due to the high allele frequency of the rs11231341 encoding the p.Tyr501Ter. Surprisingly, the premature stop codon has not been found in primates[71] (61 chimpanzee, 10 bonobo, 43 gorilla, 9 orangutan). However, according to the three ancient human genomes, the Altai Neandertal, Denisova, Ust-Ishim and Vindija, the ancient humans were either p.Tyr501Ter homozygous or heterozygous (available from the ancient genome browser, https://bioinf.eva.mpg.de/jbrowse/)[72, 73]. These data suggest that the mutation was introduced in the ancient human period, the reason for which remains an open evolutionary question.

Examination of GWAS and publicly available databases (GWAS Catalog, NCBI Phenotype-Genotype Integrator) revealed strong associations between SNPs in SLC22A24 and plasma levels of various metabolites, particularly those of steroid glucuronide conjugates (S1 Table), as well as steroid sulfate conjugates (S1 Table[25]). SLC22A24 p.Tyr501Ter (rs11231341), which has no transport activity and no protein expression on the plasma membrane (Fig 6), is associated with lower plasma levels of steroid glucuronides in three independent cohorts and in multiple ethnic groups (Table 1). Thus, it is reasonable to speculate that individuals harboring p.Tyr501Ter would have reduced plasma steroid glucuronide levels due to their reduced ability to reabsorb these metabolites. Interestingly, we observed that SLC22A24 possesses a PDZ-consensus motif at its C-terminal, which is also present in several known transporters localized to the apical membrane of the renal proximal tubule [74, 75] (S11 Table). The presence of this motif supports the direction of the observed associations and the idea that the transporter functions in the reabsorption of steroid conjugates in the kidney. Although it has been shown that the ABC transporters in the proximal tubule, MRP2 and MRP3 function in the efflux of a steroid conjugates into the urine and plasma, respectively[76], to our knowledge, there have been no reports characterizing renal reabsorptive SLC transporters for steroid glucuronide conjugates. Glucuronide conjugates of steroids normally undergo net secretion in the

proximal tubule, presumably by MRP2 and the organic anion transporters, OAT1 and OAT3 [76]. Though our data suggest that SLC22A24 functions in the reabsorption of steroid metabolites in the kidney, future studies are needed to confirm its role. In particular, measurements of steroid metabolites in both urine and plasma are required to understand the role of the transporter in their renal clearance. Unfortunately, urinary levels of steroid conjugates were not available in ARIC, which was used as our discovery cohort.

The highly significant association between the SLC22A24 nonsense variant and androsterone glucuronide levels led us to examine associations between the variant and diseases that are known to be associated with this metabolite. Through a phenome-wide scan of multiple publicly available sources, we observed associations between SLC22A24 and conotruncal heart defect ($p = 1x10^{-6}$)[77], lipid levels, and cardiovascular disease (p-values range from $1x10^{-5}$ to $9x10^{-5}$, data source Phenotype-Genotype Integrator, https://www.ncbi.nlm.nih.gov/gap/phegeni). In addition, we used phenome-wide association study (PheWAS) approach to analyze the many phenotypes in UKBiobank (http://geneatlas.roslin.ed.ac.uk/phewas/) with SLC22A24 SNPs that are associated with steroid and steroid conjugate levels, rs11231341 (p. Tyr501Ter) and rs78176967 (S12 Table). Hematological measurements (e.g. immature reticulocyte fraction, neutrophil percentage) are among the top traits that are significantly associated with the two SNPs (S12 Table). These traits have been previously associated with steroid hormones [78–81]. Further, studies have shown that levels of androsterone glucuronide are elevated in women with acne and/or hirsutism [82–87]. Through a detailed search of associations with acne and/or hirsutism, we noted that the top SNPs in the SLC22A24 locus associated with acne ($p = 0.001$) in the UKBiobank[88] are among the top SNPs significantly associated with steroid and steroid glucuronide levels (S13 Table). Overall, the identified phenotypic data is consistent with the idea that dysregulation of steroid hormone metabolism contributes to the presentation of acne, blood traits, and cardiovascular disease in humans (S12 and S13 Tables).

In conclusion, these studies represent the first functional characterization of SLC22A24, a transporter in the UST (Unknown Substrate Transporter) region of the genome. The transporter has a unique substrate specificity and our studies suggest that it functions in the reabsorption of steroid glucuronide conjugates in the kidney (Fig 7C). This work represents an important step in the global effort to functionally characterize the remaining members of the SLC22 family.

Materials and methods

Sources used to determine associations between genetic variations in SLC22A24 and human traits and diseases

Various public sources were used to obtain significant associations of SNPs within SLC22A24 gene +/- 50,000 kb (chr11:62553988–62743269 (hg18); chr11: 62797412–62986693 (hg19); chr11:63029940–63219221 (hg38)) with clinical variables related to metabolites levels. These sources included the GWAS Catalog, Phenotype-Genotype Integrator, GRASPS, GWAS Metabolic Server (http://metabolomics.gwas.eu) and others. The results from the search in public sources are shown in S1 Table.

Further association analyses of steroid metabolite levels in the Atherosclerosis Risk in Communities Study (ARIC)

An application was approved by the ARIC Publication Committee to determine the associations of *SLC22A24* variants with steroid and bile acid metabolites in ARIC cohort (Manuscript proposal #3167). Below we described the ARIC study population used in this association

analyses, metabolite measurement and association analyses of selected SNPs and metabolites relevant to this study. *Study population*: The ARIC is a prospective epidemiologic study conducted in U.S., designed to investigate the etiology and natural history of atherosclerosis, the etiology of clinical atherosclerotic diseases, and variation in cardiovascular risk factors, medical care and disease by race, gender, location, and date[89]. Using this cohort, investigators have published several papers related to associations between serum metabolome and diseases or traits among the subjects in the ARIC study[90]. In general, each visit included interviews and a physical examination [89]. In 2014, Metabolomic profiles were measured in baseline serum from 599 African- and 1,553 European-Americans. Participants were excluded if they did not give consent for use of DNA information. *Assessment of Metabolomic Profiles*: Methods for the assessment of metabolomics profiles, and whole exome sequencing were described previously [91–93]. In brief, metabolites were completed by Metabolon Inc. (Durham, USA) using an untargeted, gas chromatography-mass spectrometry and liquid chromatography-mass spectrometry (GC-MS and LC-MS)-based metabolomic quantification protocol. The annotated exome was captured by NimbleGen's VCRome2.1 (Roche NimbleGen), and the captured exons were sequenced using Illumina HiSeq 2000. A total of 573 African Americans and 1375 Caucasians were available for the association analyses. The primary analysis was to determine the association of rs11231341 (SLC22A24 p.Tyr501Ter) with androsterone glucuronide (X-22379) and etiocholanolone glucuronide.

Association of steroid glucuronide in Middle Eastern population

Yousri NA et al. have published the whole exome sequencing to identify genetic polymorphisms associated with metabolite levels in Middle Eastern population[9]. Using the data computed from Yousri NA et al.[9], the beta and p-value for the association of SLC22A24 p.Tyr501Ter (rs11231341) with androsterone glucuronide (X-22379) and etiocholanolone glucuronide are shown in Table 1.

Establishment of transient and stable cell lines

Human embryonic kidney cell lines (HEK293) containing a Flp-In expression vector (HEK293 Flp-In) were used to create either transient or stable cell lines expressing SLC22A24 cDNA or vector only DNA. SLC22A24 (NM_173586 and NM_001136506) tagged open reading frame (ORF) clone and vector control (pCMV6-Entry) were purchased from OriGene. These vectors were either transiently transfected or stably transfected into Flp-In 293 cells using Lipofectamine LTX (Thermo Fisher Scientific). 500 ng of DNA and 1 μL of Lipofectamine LTX or 10 μg of DNA and 44 μL of Lipofectamine LTX were used for transfection into each well of the 24-well plate (seeding density $2.0x10^5$ cells/well) or 100 mm tissue culture plate (seeding density $4x10^6$ cells/well) respectively. More detailed methods to create stably or transiently transfected cells have been described from our groups (see [26, 38, 94]). For transient transfection, after 36 to 48 hours, cells were use for transporter studies (see section called Transporter uptake studies) or for protein quantifications (see section called Analysis of SLC22A24 protein levels in transiently expressing HEK293 cells.). For stable cell lines, after 48 hours, cells were transferred to a new 100 tissue culture plate and were treated with 500 μg/mL Geneticin. Fresh media containing 500 μg/mL Geneticin was replaced each day.

Cloning and expression profiling of SLC22A24 in liver and kidney

For the cloning, pooled total RNA of human kidney and liver were purchased from Clontech. Total RNA (2 μg) from each sample was reverse transcribed into cDNA using SuperScript VILO cDNA Synthesis kit (Thermo Fisher Scientific) according to the manufacturer's

protocol. For the expression profiling of SLC22A24, the cDNA from pooled human tissues was purchased from Takara Bio Inc. In particular the cDNA from two tissue panels was purchased, Human MTC Panel I and II. S7 Table shows the primers that were used in the PCR to clone the transcripts, NM_173586 and NM_001136506. PCR products were cloned into pcDNA5FRT and sequenced (MCLAB, South San Francisco) to confirm the transcript. In addition, the primers that were used in the PCR to detect the presence of GAPDH (housekeeping gene) and the three SLC22A24 isoforms are shown in S7 Table. For expression profiling study, 10 ng cDNA of various tissues (Takara Human MTC panel I & II) were used for each reaction and amplified by KOD Xtreme Hot Start DNA polymerase kit (Takara). The PCR cycling conditions used are: (i) activation at 94˚C for 2 min, (ii) denature at 98˚C for 10 sec, (iii) annealing at 57.5˚C for 30 sec and (iv) extension at 68˚C for 1 min.

RNA isolation and quantitative RT-PCR

HEK293 Flp-In cells were grown in 24-well poly-D-lysine coated plates (seeding density of 1.5–1.8 x10^5/well) until 75–80% confluency. Upon reaching confluency, cells were transiently transfected with vector only or vector containing different SLC22A24 isoforms (in pCMV6-Entry expression vector). The different isoforms were: SLC22A24 with 552 amino acids (NM_001136506.2, ENST00000612278.4), SLC22A24-551 with 551 amino acids (NM_001136506.2, ENST00000417740.5) and SLC22A24-322 with 322 amino acids (NM_173586.2) and SLC22A24 p.Tyr501Ter (c.1503T>G) (in both pCMV6-Entry and pcDNA5FRT vector backbone). 500 ng of the plasmid DNA, 1 μL of Lipofectamine LTX (Thermo Fisher Scientific) and 100 μL of the Opti-MEM I reduced serum media (Thermo Fisher Scientific) were used in the transfection mixture. 40 hours after transfection, media was removed and 350 μL of RNA Lysis buffer was added to each well. Total RNA was isolated using the Qiagen RNeasy kit (Qiagen). cDNA was synthesized using SuperScript VILO cDNA Synthesis Kit (Thermo Fisher Scientific). Quantitative RT-PCR (qRT-PCR) was performed using Taqman reagents and specific primer and probe sets for human SLC22A24 (Assay ID: Hs00543210_m) and GAPDH (Assay ID: Hs99999905_m1) (Applied Biosystems, Foster City, CA). Reactions were performed in a 96-well plate with a 10 μL reaction volume using an ABI 7900HT Fast Real-time PCR system (Applied Biosystems) using the default instrument settings. Expression levels were determined from three independent biological samples using the Ct method after normalization to endogenous levels of GAPDH[56, 95]. The results are expressed as a fold increase of the *SLC22A24* mRNA level over the cell lines expressing the vector control.

Synthesis of *SLC22A24* ortholog cDNA

Sequences of SLC22A24 orthologs from mouse lemur (*Microcebus murinus*, Transcript ID ENSMICT00000042921.1), rabbit (*Oryctolagus cuniculus*, XM_002720992.3, Transcript ID ENSOCUT00000010486.2), and horse (*Equus caballus*, Transcript ID ENSECAT00000014835.1) were obtained from ensembl.org (release 95)[45] and the National Center for Biotechnology Information (NCBI). The cDNA of the gene was synthesized by GeneScript and inserted into the BamHI and XhoI restriction sites on the expression vector pcDNA5FRT (Invitrogen). The cDNA of the rat Slc22a9/Slc22a24 (NM_173302.1) was purchased from OriGene (Catalog number: RR202601). The synthesized *SLC22A24* ortholog cDNA clones were used to make GFP fusion constructs to determine their subcellular localization. The GFP coding sequence was ligated to the 3' end of the SLC22A24 ortholog cDNA in the expression vector pcDNA5/FRT using In-Fusion HD cloning kit (ClonTech #639642). The human SLC22A24 GFP fusion construct was purchased from OriGene (Catalog number: RG227944). The primers used to clone

the GFP fusion constructs for the other non-human species are shown in S9 Table. The sequence of the constructs was confirmed by sequencing (MCLAB, South San Francisco).

Transporter uptake studies

HEK293 Flp-In cells stably or transiently transfected with *SLC22A24* were seeded at a density of 180,000 to 200,000 cells/0.5 mL in poly-D-lysine 24-well plates approximately 16 to 24 hours prior to uptake studies. For uptake studies in transiently expressing SLC22A24, methods described in previous section were performed before this study. Before the uptake studies, the culture medium (Dulbecco's modified Eagle's medium, DMEM H-21, supplemented with 10% fetal bovine serum) was removed and cells were incubated in 1.0 mL HBSS for 10–20 min at 37 ˚C. For the screening of radiolabeled compounds as substrate of SLC22A24, trace amount of the radiolabeled compounds (^{3}H or ^{14}C) were diluted in the HBSS (1:2000 or 1:3000) for uptake experiments. The details of the compound concentrations and uptake time are described in the results or figure legends. Uptake reactions were terminated by washing cells twice with 1 mL HBSS buffer and the cells were incubated in 700–800 μL of lysis buffer (0.1 N NaOH, 0.1% v/v SDS) and then 650–750 μL of cell lysate was transferred to scintillation fluid for scintillation counting. Experimental condition used in this study was similar to transporter assays published previously by our group[95, 96]. For pH dependence experiments, HBSS buffer was adjusted to different pHs (5.5, 7.4 and 8.5) by adding hydrochloric acid or sodium hydroxide[38]. For chloride dependence study, two different uptake buffers were used: (1) chloride free buffer (125 mM sodium gluconate, 4.8 mM potassium gluconate, 1.2 mM magnesium sulfate, 1.3 mM calcium gluconate, and 5 mM HEPES; adjusted with sodium hydroxide to pH 7.4); or (2) sodium buffer (125 mM sodium chloride, 4.8 mM potassium chloride, 1.3 mM calcium chloride, 1.3 mM magnesium sulfate, and 5 mM HEPES, adjusted with sodium hydroxide to pH 7.4)[95, 97]. For *trans*-stimulation studies, experimental conditions described in our previously published methods were used[38]. In brief, the SLC22A24 or EV stable cell lines were pre-incubated with either buffer or 2 mM succinic acid, 2 mM α-ketoglutaric acid or 2 mM glutaric acid for 2 hours. Then, the cells were washed twice with HBSS before initiating uptake of the anions (estrone sulfate, estradiol glucuronide, androstanediol glucuronide or taurocholic acid).

Kinetic studies of steroid conjugates and bile acid transport by SLC22A24

We characterized the kinetics of the three steroid conjugates and two bile acids by SLC22A24, discovered from the initial screening. We selected SLC22A8 for comparison to SLC22A24 due to the significant accumulation of steroid conjugates by SLC22A8. The kinetic studies were performed following methods developed in our group[38]. In brief, these studies were performed using stable cell lines expressing SLC22A24 (NM_001136506) or SLC22A8[98]. For the kinetic studies of bile acid, only SLC22A24 expressing cells were used, as bile acids are poor substrates of human SLC22A8[99]. For each substrate, we varied the concentrations of the unlabeled compounds. The uptake rate was evaluated at 1 minute, as this fell within the linear range in the uptake versus time plot for each substrate. Each point represents the mean ± SD uptake in the transporter-transfected cells minus that in empty vector cells. The plots show the result for a representative experiment of n = 2.

Inhibition of SLC22A24-mediated uptake by various anions and drugs

HEK293 Flp-In cells stably transfected with SLC22A24 (NM_001136506) were seeded at a density of 200,000 to 220,000 cells/0.5 mL in poly-D-lysine 24-well plates approximately 16 to 24 hours prior to the experiments. On the day of the experiments, cells were washed with 1 mL

of Hank's buffered salt solution (HBSS) per well and then preincubated for 10–20 min in the 1 mL of HBSS buffer. To assess inhibition, the cells were incubated with uptake buffer (1 μM unlabeled taurocholic acid and trace amount of [^{3}H]-taurocholic acid in HBSS) containing either 100 μM estrone sulfate as positive control, 1% DMSO as negative control or 100 μM of the selected test compounds. Each of the 24-well plates contained negative and positive controls. Uptake of [^{3}H]-taurocholic acid was stopped after 10 min by washing twice with ice-cold HBSS. The cells were lysed in 900 μL of lysis buffer (0.1 N NaOH and 0.1% SDS in double distilled water) per well while shaking for up to 90 min. 820 μL of the lysate was then added to 2.5 mL of EcoLite scintillation fluid (MP Bio), and the radioactivity was determined on a LS6500 Scintillation Counter (Beckman Coulter). Values were corrected for protein concentration as determined with a BCA assay kit (Thermo Scientific, Rockford, IL). All values were determined in triplicate, and final values are expressed as % uptake relative to negative control (1% DMSO). Methods used in the inhibition studies were from previous studies in our laboratory [100, 101]. The test compounds were tested at least one time and some compounds were selected to test two or three times to check for consistent results.

Analysis of SLC22A24 protein levels in transiently transfected HEK293 cells

A 100 mm tissue culture plate of HEK293 Flp-In cells was transiently transfected with the following expression vectors: vector only (pCMV6-Entry), SLC22A24 (NM_001136506), SLC22A24-322 (NM_173586), SLC22A24 p.Tyr501Ter (created using site directed mutagenesis). All proteins contained a C-terminal Myc-DDK tag. For evaluation of the SLC22A24 protein levels, cell lysates were collected 48 hours after transfection. Plasma membranes were separated from intracellular membrane using the Plasma Membrane Protein Extraction Kit (Abcam ab65400)[32]. Transfected cells were harvested in 2 mL of homogenization buffer. Cells were homogenized using the syringe-based homogenization method. A needle (27 gauge) was attached to a 1 mL syringe. The cells lysates were passed through the syringe 5 times. A higher gauge needle, 30 gauge, was then used and the suspension was passed through for 5 additional times. The plasma membrane fraction was resuspended in 50 μL lysis buffer. Cell samples were deglycosylated using PNGase F (NEB P0708L) per the manufacturer's protocol.

Plasma membrane localization studies

The resulting GFP fusion constructs were transiently transfected in human embryonic kidney 293 (HEK293) cells. The cells were maintained in Dulbecco's modified Eagle's Medium (DMEM) (Fisher) supplemented with 10% fetal bovine serum (FBS) (Invitrogen) and 10 mM HEPES (Fisher). Dr. R. Irannejad's laboratory has established standard methods for transfecting HEK293 cells for membrane localization studies[102–104]. HEK293 cells were grown to ~70% confluence in poly-L-lysine coated glass coverslips in 12-well plates. The cells were transiently transfected with 1 μg plasmid DNA with 2 μL of Lipofectamine 2000 (Thermo Fisher Scientific). 24-hours after transient transfection, cells were fixed in 4% paraformaldehyde. Coverslips were mounted using ProLong Gold Antifade Mountant (Thermo Fisher Scientific) on glass microscope slides and visualized using Yokogawa CSU-22 Spining Disk Confocal microscope (Nikon Instruments, Inc).

Metabolomic study

Samples of cell pellets from HEK293 Flp-In were prepared and sent to Metabolon Inc. (North Carolina, USA) for metabolites measurement using Metabolon platform. Cell pellet samples

were prepared from five 100 mm culture dishes each of HEK293 Flp-In stably expressing vector only (pCMV6-Entry) or cells transiently expressing SLC22A24 cDNA (NM_001136506). 24 hours after transient transfection, media were removed and DMEM-H21, phenol red free, media (Invitrogen) containing 20% Fetal Bovine Serum, U.S.D.A. Approved Origin (Product Number 89510–186, Lot Number 356B16) were added to each dish. 48 hours after transient transfection, the cells were washed twice with cold PBS, and cells were scraped and transferred to a 15 mL falcon tubes. Tubes were centrifuged at 3000 rpm and supernatant were removed and stored the cell pellet in -80 ˚C freezer until shipping to Metabolon Inc. for metabolomics measurement (see[26] to obtain methods for metabolomic study).

Metabolomic data analysis

Raw metabolomic data were processed by Metabolon[105]. Individual metabolites were quantified and the values normalized. A two-tailed t-test was performed between EV and SLC22A24 and between EV and SLC22A1. False discovery rate (FDR) was used to correct the p-values. We eliminated potential false positive metabolites using the following criteria: if FDR < 0.05 between SLC22A24 and EV cells, but p-value > 0.05 between SLC22A1 and EV cells, the metabolites were considered false positives. The results of the analysis are included in S3 Table.

Clustering of human SLC22A24 and other species orthologs in the SLC22 family

Three phylogenetic trees were constructed. The first consists of the UniProt amino acid sequences for the human SLC22 family members with known substrates, human SLC22 members in the gene cluster between SLC22A8 and SLC22A9 of the human chromosome 11q12.3 (hg19 assembly), and all identified mouse and rat SLC22 members (Fig 5A). The second and third tree were built focusing on anion transporters. The second tree (Fig 5B) includes the human SLC22A24, known SLC22A24 orthologs in other species identified in the UniProt, known human anion transporters in SLC22 family (SLC22A6, A7, A8, A9, A11 and A12), genes in the human chromosome 11q12.3 (SLC22A10, A25), mouse anion transporters (Slc22a6, a7, a8, a9/a19, a9/a21, a12), mouse Slc22 genes in the clusters between Slc22a8 and Slc22a19 of the mouse chromosome 19qA (Slc22a26, a27, a28, a29, a30), rat anion transporters (Slc22a6, a7, a8, a9/a24, a12) and Slc22 genes in the clusters between Slc22a8 and Slc22a19 of the rat chromosome chr1q43 (Ust4r, Ust5r, Slc22a25). The third tree (S5 Fig) includes human SLC22A24 and the reference proteomes of all chordates (https://www.uniprot.org/help/reference_proteome), which have sequence similarity to human SLC22A24 with an e-value $< 1x10^{-200}$. For each tree, a multiple sequence alignment was created using MUSCLE[106], with default parameters on the CIPRES Science Gateway [107]. Phylogenetic trees were built using a RAxML[108] installation "RAxML-HPC2 on XSEDE" on the CIPRES Science Gateway. The tree was calculated on protein sequences using the protein GAMMA model with the JTT substitution matrix and rapid bootstrapping with 100 replicas. A resulting extended majority rule consensus tree was visualized using Dendroscope 3[109]. See S1 and S2 Appendices for the amino sequences used to create the phylogenetic tree in Fig 5A and 5B and S5 Fig.

Comparative structure modelling

Due to the lack of available crystal structures, human OCT1 (SLC22A1), OCTN1 (SLC22A4), OAT1 (SLC22A6), and SLC22A24 were modeled using the crystal structure of the maltose-bound human GLUT3 (SLC2A3) in the outward-open conformation at 2.6 Å as a template

(PDB ID 4ZWC)[110]. The template was selected on the basis of the shared MFS fold assignment [10], structure quality, and sequence similarity to the targets of interest (40%, 38%, 39% sequence similarity to SLC22A1, SLC22A6, SLC22A24, respectively). The sequence alignment was obtained by a manual refinement of gaps in the output from the PROMALS3D server (S1A Fig) [111]. The maltose molecule from the crystal structure was treated as the BLK residue type in the alignment and was copied from the template structure into a model as a rigid body. One hundred models were generated for each protein using the "automodel" class of MODELLER 9.16 (https://salilab.org/modeller/). The models had acceptable normalized discrete optimized protein energy scores (zDOPE)[112, 113]. The top scoring models were selected for electrostatic potential analysis with the APBS electrostatics plugin[114] within PyMOL 2.0.3 using the AMBER force field with default settings.

Site directed mutagenesis to create SLC22A24 p.Tyr501Ter

We used the Gibson Assembly protocol from New England BioLabs (catalog no. E5510) to create the SLC22A24 p.Tyr501Ter construct by amplifying unique segments from the SLC22A24 wild type and pCMV6-Entry vectors. The following primers were used to amplify the segment containing SLC22A24 residues 1–500 from the SLC22A24 wild type plasmid:

forward primer GCCGCCGCGATCGCCATGGGCTTTGATGTGCTCCT

reverse primer CGGCCGCGTACGCGTGGAAATCCAGGGTAGGTGGG.

The following primers were used to amplify the segment containing the linker, tag, stop codon, and the rest of the vector backbone from the pCMV6-Entry vector:

forward primer CTACCCTGGATTTCCACGCGTACGCGGCCGCTCGA

reverse primer CACATCAAAGCCCATGGCGATCGCGGCGGCAGATC

Each segment contains a 15 bp overhang necessary for the Gibson Assembly workflow. The amplified segments were then assembled following the manufacturer's protocol (https://www.neb.com/protocols/2012/12/11/gibson-assembly-protocol-e5510).

Kidney tissue procurement

Frozen human postmortem frozen renal cortical tissues from donors aged 3 to 29 years old were obtained from the NIH NeuroBioBank at the University of Maryland, Baltimore, MD [115, 116]. These tissues were selected from patients, who are of self-reported African American descent. Tissues were procured at the time of autopsy within 48 hours after death and were kept frozen at -80˚C for later research. Tissues were selected for having no renal abnormalities in pathology and primary diagnosis. Twenty kidney samples were selected for quantitative proteomic analysis (see section below).

Genotype kidney tissues for SLC22A24 c.1503T>G

Genomic DNA from the 20 renal cortical tissues procured from the NIH NeuroBioBank were isolated using Wizard genomic DNA purification kit (Promega). The following forward (CCACAAGGGCAGAAAGTATG) and reverse primers (CAAATTATCAAAGCTGCGGG TTA) were used to PCR the genomic region of SLC22A24 and using Sanger sequencing primer (TAAGCCAGATATTGTTCACG) to determine the genotype for the stop-gained variant in SLC22A24, c.1503T>G (rs11231341).

Quantitative proteomics

Absolute quantitative proteomics was performed on the 20 renal cortical tissues. SLC22A24 and SLC22A24-551 were quantified using LC-MS/MS with Selective Reaction Monitoring (SRM) approach by Omics Technologies, Inc. (http://www.proteomics.omicstech.com/proteomics/srm.html, previously known as MyOmicsDx, Inc., Towson, MD, USA). The following are the main steps for quantitative proteomic analysis for the two transcript isoforms of SLC22A24. The detailed material and method used for each steps are described below as well as in our previous submitted manuscript[116]. The material and methods described below were provided by Omics Technologies, Inc.

- Membrane protein extraction and protein preparation
- Chemicals and Reagents
- Preparation of the Chromatography Solutions
- SRM analysis
- Tuning of the Agilent 6495 Triple Quadrupole mass spectrometer
- SRM data analysis

Membrane protein extraction and protein preparation. Frozen renal cortical tissues were processed to extract membrane proteins. Membrane protein samples were then processed by Omics Technologies, Inc. using "Filter Assisted Sample Preparation" (FASP) method [117]. Briefly, protein samples in 9M UREA were reduced with 5 mM TCEP at 37 ˚C for 45 min and reduced cysteine were blocked using 50 mM iodoacetmide (IAA) at 25 ˚C for 15 min. Protein samples were then cleaned using 10 kDa Amicon Filter (UFC501096, Millipore) three times using 9M urea and two times using MyPro-Buffer 1 (Omics Technologies, Inc.). Samples were then proteolyzed with trypsin (V5111, Promega) for 12 hrs at 37 ˚C. The peptide solution then was acidified by adding 1% trifluoroacetic acid (TFA) and was incubated at room temperature for 15 min. A Sep-Pak light C18 cartridge (Waters Corporation) was activated by loading 5 mL 100% (vol/vol) acetonitrile and was washed by 3.5 mL 0.1% TFA solution two times. Acidified digested peptide solution was centrifuged at 1,800 × g for 5 min, and the supernatant was loaded into the cartridge. To desalt the peptides bound to the cartridge, 1 mL, 3 mL, and 4 mL of 0.1% TFA were used sequentially. To elute the peptides from the cartridge, 2 mL of 40% (vol/vol) acetonitrile with 0.1% TFA was used. The eluted peptides were lyophilized overnight and reconstituted in 37 μL MyPro-Buffer 3 (Omics Technologies, Inc.).

Chemicals and reagents. TCEP (TCEP (tris-(2-carboxyethyl) phosphine) and MMTS (methyl methanethiosuphonate) were purchased from Thermo Scientific (Waltham, Massachusetts). LysC and Trypsin proteases were purchased from Promega (Fitchburg, Wisconsin). C18 Cartridges for sample preparation, and chromatography columns for basic reverse phase liquid chromatographic (bRPLC) and online HPLC of Triple Quadrupole mass spectrometer were purchased from Waters (Milford, Massachusetts). Acetonitrile was purchased from JT Baker, and formic acid was obtained from EMD Millipore (Billerica, MA, USA). MyPro-Buffer 1, MyPro-Buffer 2 and MyPro-Buffer 3 were utilized by Omics Technologies, Inc. All other reagents were purchased from Sigma-Aldrich (St. Louis, Missouri) unless otherwise indicated.

Preparation of the chromatography solutions. bRPLC solvent A contained 10mM triethylammonium bicarbonate buffer (TEABC); bRPLC solvent B contained 10mM TEABC, 90% (vol/vol) acetonitrile. SRM Mass Spectrometry solvent A was comprised of water with 0.1% (vol/vol) formic acid; SRM solvent B was acetonitrile with 0.1% (vol/vol) formic Acid.

SRM analysis. Peptide samples reconstituted in 37 μl MyPro-Buffer 3 (MyOmicsDx, Inc) were spiked with MyPro-SRM Internal Control Mixture (MyOmicsDx, Inc) composed of a pool of 1 femto mole heavy isotope labeled peptides covering a large hydrophobicity window and a large M/z range (M/z 200 ~ 1300), and were subject to SRM analysis. Peptide samples were eluted through an online Agilent 1290 HPLC system into the Jet Stream ESI source of an Agilent 6495 Triple Quadrupole Mass spectrometer.

Tuning of the Agilent 6495 Triple Quadrupole mass spectrometer. After every preventative maintenance, Agilent 6495 Triple Quadrupole mass spectrometer was tune using manufacturer's tuning mixture followed by MyPro-SRM Tuning Booster (MyOmicsDx SRM tuning mixture). Before and after every batch of SRM analysis, to ensure the stable and consistent performance of the mass spectrometer throughout the entire study, MyPro-SRM Performance Standard (MyOmicsDx), a mixture of standard peptides across a wide range of mass (M/z 100–1400) and a broad range of hydrophobicity were analyzed.

SRM data analysis. The abundance of a target peptide was represented by the total area under the curve (AUC) of all its transitions normalized to the total AUC of all transitions from the most nearby (sharing a similar hydrophobicity) heavy isotope-labeled peptide from MyPro-SRM Internal Control Mixture (MyOmicsDx, Inc) spiked in before the SRM analysis. Absolute quantification of each protein is performed through applying AQUA Peptides purchased from Sigma-Aldrich. A summary of all quantification results were shown in S8 Table.

Supporting information

S1 Fig. A common splice donor variant, rs1939749, found in 3'-end of exon 1. (A) A splice donor, rs1939749, C>T (+ DNA strand), is strongly associated with lower expression levels of SLC22A24 in kidney tubular cells observed in two independent kidney eQTL databases, http://susztaklab.com/eqtl and http://nephqtl.org/, p-values $3.7x10^{-23}$ (beta = 1.2, N = 121) and $1.9x10^{-13}$ (beta = 0.79, N = 166) respectively. (B) Zoom in of the SLC22A24 gene region (hg19) flanking exon 1 and exon 2 with padding of 25 bases. This splice donor, rs1939749, C>T (+ DNA strand), creates a premature stop.
(TIF)

S2 Fig. Steroid conjugates and bile acids uptake in HEK293 cells expressing SLC22A24 and SLC22A8 as a function of time. (Left) Uptake of [^{3}H]-estrone sulfate (with 1 μM unlabeled estrone sulfate) as a function of time in cells expressing SLC22A24 and SLC22A8. The uptake rates for estrone sulfate and two other steroid glucuronides, estradiol glucuronide and androstanediol glucuronide were evaluated at 1 min. (Middle and right) Uptake of two bile acids, [^{14}C]-glycocholic acid (with 1 μM unlabeled glycocholic acid) and [^{3}H]-taurocholic acid (with 1 μM unlabeled taurocholic acid) as a function of time in cells expressing SLC22A24 and SLC22A8 (only glycocholic acid). The uptake rate for glycocholic acid and taurocholic acid were evaluated at 1 min and 2 min respectively. Each point represents the uptake in SLC22A24- or SLC22A8-expressing cells minus that in empty vector cells. Underlying data are provided in S1 Data.
(TIF)

S3 Fig. Kinetic uptake of androstanediol glucuronide for SLC22A24. The uptake rate was evaluated at 2 minutes. Each point represents the mean ± S.D. uptake in the SLC22A24-stably transfected cells minus that in empty vector cells. Figure shows a representative plot from one experiment. The kinetic parameters of androstanediol-3-glucuronide could not be determined accurately due to a solubility limit of 300 μM. Using the plateau at 300 μM as an approximation, the affinity of androstanediol glucuronide was calculated as 741 ± 253 μM. The data were

fit to a Michaelis–Menten equation. Underlying data are provided in S1 Data.
(TIF)

S4 Fig. Effect of sodium and pH on SLC22A24-mediated uptake of anion. (A) [^{3}H]-Estrone sulfate (with 1 μM unlabeled estrone sulfate) was incubated in the Na+ buffer (125 mM sodium chloride, 4.8 mM potassium chloride, 1.3 mM calcium chloride, 1.3 mM magnesium sulfate, and 5 mM HEPES, adjusted with sodium hydroxide to pH 7.4) or Na+ free buffer (125 mM choline chloride, 4.8 mM potassium chloride, 1.3 mM calcium chloride, 1.3 mM magnesium sulfate, and 5 mM HEPES, adjusted with TRIS-HCl to pH 7.4) for 10 min. Data shown are from the mean ± S.D. for representative experiments of n = 2. (B) [^{3}H]-Estrone sulfate, [^{3}H]-estradiol glucuronide or [^{3}H]-taurocholic acid (with 1 μM unlabeled compound) was incubated in the HBSS buffer adjusted to pH 5.5, 7.4 and 8.5, and the cells were incubated for 10 min. Data shown are from the mean ± S.D. for representative experiments of n = 2. Underlying data are provided in S1 Data.
(TIF)

S5 Fig. Phylogenetic tree indicating protein sequence relationships between human USTs and identified mammalian orthologs. The unique and stable entry identifiers from UniProtKB (https://www.uniprot.org/) are shown in parentheses. Red arrow points to the human SLC22A24.
(TIF)

S6 Fig. Plasma membrane expressions of SLC22A24-GFP tagged from five different species. GFP fusion constructs were generated for the human, rabbit, horse, and mouse lemur SLC22A24 and the rat Slc22a9/Slc22a24. HEK293 cells were transiently transfected with the GFP fusion constructs. The cells were visualized by confocal microscopy to determine whether the GFP fused protein is expressed on plasma membrane.
(TIF)

S7 Fig. Plasma membrane expression, transport uptake, and predicted structure of human SLC22A24 or SLC22A24-322. (A) GFP fusion constructs were generated for human SLC22A24 and SLC22A24-322. HEK293T cells were transient transfected with the GFP fusion constructs. The cells were visualized by confocal microscopy to determine whether the GFP fused protein is expressed on plasma membrane. (B) The uptake experiments of various anions were performed for 10 minutes in triplicate wells using trace amount of the radiolabeled compound. Experiments were repeated once. Underlying data are provided in S1 Data. (C) A comparative structure model of the human SLC22A24 with 12 transmembrane helices (white and red). SLC22A24-322 is predicted to consist of transmembrane helices 1–6 (white), but missing transmembrane helices 7–12 (red).
(TIF)

S8 Fig. Schematic figure showing the amino acid differences between SLC22A24 and its isoform, SLC22A24-551. (A) Zoom in of the SLC22A24 gene region (reverse direction) between exon 9 and exon 10 along with padding of 10 bases using UCSC browser. The allele "CT" from position 1 and position 2 resulted in formation of two different isoforms (splice variants) of SLC22A24. (B) Zoom in of the SLC22A24 amino acids between exon 9 and exon 10. Arrows show the differences in amino acid positions 533 and 534, which lead to SLC22A24-551 and SLC22A24 proteins.
(TIF)

S9 Fig. Protein and transcript levels of SLC22A24 quantified by western blots and qRT-PCR. (A) Western blot of cells expressing empty vector, SLC22A24, SLC22A24 p.Tyr501Ter,

and SLC22A24-322. The latter two were not expressed on the plasma membrane despite detectable mRNA levels. (B) Transcript levels of mRNA measured using Taqman assay in HEK293 Flp-In cells transiently expressing empty vector, SLC22A24, SLC22A24-322 and SLC22A24 p.Tyr501Ter. Underlying data are provided in S1 Data.
(TIF)

S10 Fig. Alignment of SLC22A24 and SLC22A24-551. The peptides of interest and their corresponding fragments used to quantify the protein levels of SLC22A24 and SLC22A24-551 are outlined in black. Peptide 1 corresponds to the SLC22A24 and SLC22A24-551 proteins. Peptides 2 and 3 correspond to SLC22A24 and SLC22A24-551, respectively.
(TIF)

S1 Table. Metabolites that are significantly associated with SNPs within SLC22A24 locus.
(XLSX)

S2 Table. Association of SLC22A10 and SLC22A25 missense and nonsense variants with two steroid conjugates.
(XLSX)

S3 Table. Results from metabolomic study using Metabolon Inc. technology.
(XLSX)

S4 Table. Experimental data of uptake in members of the SLC22 family used as positive controls.
(XLSX)

S5 Table. Kinetic parameters of steroid conjugates and bile acids uptake by SLC22A24.
(DOCX)

S6 Table. List of metabolites or prescription drugs used in the inhibition studies.
(XLSX)

S7 Table. Source of cDNA.
(DOCX)

S8 Table. Results from SRM analysis.
(XLSX)

S9 Table. Primers and PCR conditions used in this study.
(DOCX)

S10 Table. Transcript levels of 22 genes in SLC22 family in kidney samples.
(XLSX)

S11 Table. List of renal transporters in the solute carrier (SLC) superfamily that are known to have apical membrane localization.
(DOCX)

S12 Table. Phenome-wide association results of two SNPs in SLC22A24.
(XLSX)

S13 Table. Association results of variants in SLC22A24 locus with acne/acne vulgaris extracted from UKBiobank.
(XLSX)

S1 Appendix. FASTA file containing amino acid sequences of solute carrier family 22 (human, mouse and rat). These sequences in FASTA format is used to create phylogenetic

tree in Fig 5A.
(TXT)

S2 Appendix. FASTA file containing amino acid sequences of anion transporters of the solute carrier family 22 (human, primates, mouse, rat and other species). These sequences in FASTA format is used to create phylogenetic tree in Fig 5B and S5 Fig.
(TXT)

S1 Data. Contains underlying data for figures.
(XLSX)

Acknowledgments

The authors thank the National Institute of Health (NIH) NeuroBioBank at the University of Maryland, Baltimore, M.D. for providing the frozen human kidney tissue. The authors also thank the Omics Technologies Inc. for their services to quantify the SLC22A24 protein levels in the renal cortical tissues, Dr. Tao Long, for personal communications regarding the interpretation of the GWAS metabolomic data[8], Dr. Eyal Akiva for discussions regarding the phylogenetic analysis, Professor Benjamin Humphreys for describing the online analyzer of the kidney single cell datasets, and Dr. Marc de Manuel Montero, for providing information about available data associated with his publication regarding primate genomic diversity (https://www.ncbi.nlm.nih.gov/bioproject/347464).

Author Contributions

Conceptualization: Sook Wah Yee, Adrian Stecula, Kathleen M. Giacomini.

Data curation: Sook Wah Yee, Adrian Stecula, Ling Zou, Elena V. Feofanova, Marjolein van Borselen, Kit Wun Kathy Cheung, Noha A. Yousri, Karsten Suhre, Jason M. Kinchen, Eric Boerwinkle, Roshanak Irannejad, Bing Yu.

Formal analysis: Sook Wah Yee, Adrian Stecula, Huan-Chieh Chien, Elena V. Feofanova, Marjolein van Borselen, Noha A. Yousri, Karsten Suhre, Eric Boerwinkle, Roshanak Irannejad, Bing Yu, Kathleen M. Giacomini.

Funding acquisition: Kathleen M. Giacomini.

Investigation: Kathleen M. Giacomini.

Methodology: Sook Wah Yee, Adrian Stecula, Huan-Chieh Chien, Ling Zou, Jason M. Kinchen.

Resources: Kathleen M. Giacomini.

Supervision: Kathleen M. Giacomini.

Validation: Sook Wah Yee, Huan-Chieh Chien.

Visualization: Sook Wah Yee, Adrian Stecula.

Writing – original draft: Sook Wah Yee, Adrian Stecula, Kathleen M. Giacomini.

Writing – review & editing: Sook Wah Yee, Adrian Stecula, Huan-Chieh Chien, Ling Zou, Elena V. Feofanova, Marjolein van Borselen, Kit Wun Kathy Cheung, Noha A. Yousri, Karsten Suhre, Jason M. Kinchen, Roshanak Irannejad, Bing Yu, Kathleen M. Giacomini.

References

1. Miller WL, Auchus RJ. The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. Endocr Rev. 2011; 32(1):81–151. Epub 2010/11/06. https://doi.org/10.1210/er.2010-0013 PMID: 21051590.
2. Chen Z, Tao S, Gao Y, Zhang J, Hu Y, Mo L, et al. Genome-wide association study of sex hormones, gonadotropins and sex hormone-binding protein in Chinese men. J Med Genet. 2013; 50(12):794–801. Epub 2013/09/21. https://doi.org/10.1136/jmedgenet-2013-101705 PMID: 24049095.
3. Coviello AD, Haring R, Wellons M, Vaidya D, Lehtimaki T, Keildson S, et al. A genome-wide association meta-analysis of circulating sex hormone-binding globulin reveals multiple Loci implicated in sex steroid hormone regulation. PLoS Genet. 2012; 8(7):e1002805. Epub 2012/07/26. https://doi.org/10.1371/journal.pgen.1002805 PMID: 22829776.
4. Ohlsson C, Wallaschofski H, Lunetta KL, Stolk L, Perry JR, Koster A, et al. Genetic determinants of serum testosterone concentrations in men. PLoS Genet. 2011; 7(10):e1002313. Epub 2011/10/15. https://doi.org/10.1371/journal.pgen.1002313 PMID: 21998597.
5. Prescott J, Thompson DJ, Kraft P, Chanock SJ, Audley T, Brown J, et al. Genome-wide association study of circulating estradiol, testosterone, and sex hormone-binding globulin in postmenopausal women. PLoS One. 2012; 7(6):e37815. Epub 2012/06/08. https://doi.org/10.1371/journal.pone.0037815 PMID: 22675492.
6. Ruth KS, Campbell PJ, Chew S, Lim EM, Hadlow N, Stuckey BG, et al. Genome-wide association study with 1000 genomes imputation identifies signals for nine sex hormone-related phenotypes. Eur J Hum Genet. 2016; 24(2):284–90. Epub 2015/05/28. https://doi.org/10.1038/ejhg.2015.102 PMID: 26014426.
7. Dudenkov TM, Ingle JN, Buzdar AU, Robson ME, Kubo M, Ibrahim-Zada I, et al. SLCO1B1 polymorphisms and plasma estrone conjugates in postmenopausal women with ER+ breast cancer: genome-wide association studies of the estrone pathway. Breast Cancer Res Treat. 2017; 164(1):189–99. Epub 2017/04/22. https://doi.org/10.1007/s10549-017-4243-3 PMID: 28429243.
8. Long T, Hicks M, Yu HC, Biggs WH, Kirkness EF, Menni C, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. Nat Genet. 2017; 49(4):568–78. Epub 2017/03/07. https://doi.org/10.1038/ng.3809 PMID: 28263315.
9. Yousri NA, Fakhro KA, Robay A, Rodriguez-Flores JL, Mohney RP, Zeriri H, et al. Whole-exome sequencing identifies common and rare variant metabolic QTLs in a Middle Eastern population. Nat Commun. 2018; 9(1):333. Epub 2018/01/25. https://doi.org/10.1038/s41467-017-01972-9 PMID: 29362361.
10. Perland E, Fredriksson R. Classification Systems of Secondary Active Transporters. Trends Pharmacol Sci. 2017; 38(3):305–15. Epub 2016/12/13. https://doi.org/10.1016/j.tips.2016.11.008 PMID: 27939446.
11. Cesar-Razquin A, Snijder B, Frappier-Brinton T, Isserlin R, Gyimesi G, Bai X, et al. A Call for Systematic Research on Solute Carriers. Cell. 2015; 162(3):478–87. Epub 2015/08/02. https://doi.org/10.1016/j.cell.2015.07.022 PMID: 26232220.
12. Lin L, Yee SW, Kim RB, Giacomini KM. SLC transporters as therapeutic targets: emerging opportunities. Nat Rev Drug Discov. 2015; 14(8):543–60. Epub 2015/06/27. https://doi.org/10.1038/nrd4626 PMID: 26111766.
13. Bai X, Moraes TF, Reithmeier RAF. Structural biology of solute carrier (SLC) membrane transport proteins. Mol Membr Biol. 2017; 34(1–2):1–32. Epub 2018/04/14. https://doi.org/10.1080/09687688.2018.1448123 PMID: 29651895.
14. Pelis RM, Wright SH. SLC22, SLC44, and SLC47 transporters—organic anion and cation transporters: molecular and cellular properties. Curr Top Membr. 2014; 73:233–61. Epub 2014/04/22. https://doi.org/10.1016/B978-0-12-800223-0.00006-2 PMID: 24745985.
15. Zhang L, Dresser MJ, Gray AT, Yost SC, Terashita S, Giacomini KM. Cloning and functional expression of a human liver organic cation transporter. Mol Pharmacol. 1997; 51(6):913–21. Epub 1997/06/01. https://doi.org/10.1124/mol.51.6.913 PMID: 9187257.
16. Gorboulev V, Ulzheimer JC, Akhoundova A, Ulzheimer-Teuber I, Karbach U, Quester S, et al. Cloning and characterization of two human polyspecific organic cation transporters. DNA Cell Biol. 1997; 16(7):871–81. Epub 1997/07/01. https://doi.org/10.1089/dna.1997.16.871 PMID: 9260930.
17. International Transporter C, Giacomini KM, Huang SM, Tweedie DJ, Benet LZ, Brouwer KL, et al. Membrane transporters in drug development. Nat Rev Drug Discov. 2010; 9(3):215–36. Epub 2010/03/02. https://doi.org/10.1038/nrd3028 PMID: 20190787.
18. Wright SH. Role of organic cation transporters in the renal handling of therapeutic agents and xenobiotics. Toxicol Appl Pharmacol. 2005; 204(3):309–19. Epub 2005/04/23. https://doi.org/10.1016/j.taap.2004.10.021 PMID: 15845420.

19. Zhu C, Nigam KB, Date RC, Bush KT, Springer SA, Saier MH Jr., et al. Evolutionary Analysis and Classification of OATs, OCTs, OCTNs, and Other SLC22 Transporters: Structure-Function Implications and Analysis of Sequence Motifs. PLoS One. 2015; 10(11):e0140569. Epub 2015/11/05. https://doi.org/10.1371/journal.pone.0140569 PMID: 26536134.

20. Koepsell H. The SLC22 family with transporters of organic cations, anions and zwitterions. Mol Aspects Med. 2013; 34(2–3):413–35. Epub 2013/03/20. https://doi.org/10.1016/j.mam.2012.10.010 PMID: 23506881.

21. Hediger MA, Clemencon B, Burrier RE, Bruford EA. The ABCs of membrane transporters in health and disease (SLC series): introduction. Mol Aspects Med. 2013; 34(2–3):95–107. Epub 2013/03/20. https://doi.org/10.1016/j.mam.2012.12.009 PMID: 23506860.

22. Li S, Sanna S, Maschio A, Busonero F, Usala G, Mulas A, et al. The GLUT9 gene is associated with serum uric acid levels in Sardinia and Chianti cohorts. PLoS Genet. 2007; 3(11):e194. Epub 2007/11/14. https://doi.org/10.1371/journal.pgen.0030194 PMID: 17997608.

23. Vitart V, Rudan I, Hayward C, Gray NK, Floyd J, Palmer CN, et al. SLC2A9 is a newly identified urate transporter influencing serum urate concentration, urate excretion and gout. Nat Genet. 2008; 40(4):437–42. Epub 2008/03/11. https://doi.org/10.1038/ng.106 PMID: 18327257.

24. Suhre K, Shin SY, Petersen AK, Mohney RP, Meredith D, Wagele B, et al. Human metabolic individuality in biomedical and pharmaceutical research. Nature. 2011; 477(7362):54–60. Epub 2011/09/03. https://doi.org/10.1038/nature10354 PMID: 21886157.

25. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, et al. An atlas of genetic influences on human blood metabolites. Nat Genet. 2014; 46(6):543–50. Epub 2014/05/13. https://doi.org/10.1038/ng.2982 PMID: 24816252.

26. Yee SW, Giacomini MM, Hsueh CH, Weitz D, Liang X, Goswami S, et al. Metabolomic and Genome-wide Association Studies Reveal Potential Endogenous Biomarkers for OATP1B1. Clin Pharmacol Ther. 2016; 100(5):524–36. Epub 2016/07/23. https://doi.org/10.1002/cpt.434 PMID: 27447836.

27. Machiela MJ, Chanock SJ. LDassoc: an online tool for interactively exploring genome-wide association study results and prioritizing variants for functional investigation. Bioinformatics. 2018; 34(5):887–9. Epub 2017/10/03. https://doi.org/10.1093/bioinformatics/btx561 PMID: 28968746.

28. Gillies CE, Putler R, Menon R, Otto E, Yasutake K, Nair V, et al. An eQTL Landscape of Kidney Tissue in Human Nephrotic Syndrome. Am J Hum Genet. 2018; 103(2):232–44. Epub 2018/07/31. https://doi.org/10.1016/j.ajhg.2018.07.004 PMID: 30057032.

29. Qiu C, Huang S, Park J, Park Y, Ko YA, Seasock MJ, et al. Renal compartment-specific genetic variation analyses identify new pathways in chronic kidney disease. Nat Med. 2018; 24(11):1721–31. Epub 2018/10/03. https://doi.org/10.1038/s41591-018-0194-4 PMID: 30275566.

30. Ko YA, Yi H, Qiu C, Huang S, Park J, Ledo N, et al. Genetic-Variation-Driven Gene-Expression Changes Highlight Genes with Important Functions for Kidney Disease. Am J Hum Genet. 2017; 100(6):940–53. Epub 2017/06/03. https://doi.org/10.1016/j.ajhg.2017.05.004 PMID: 28575649.

31. Chen L, Shu Y, Liang X, Chen EC, Yee SW, Zur AA, et al. OCT1 is a high-capacity thiamine transporter that regulates hepatic steatosis and is a target of metformin. Proc Natl Acad Sci U S A. 2014; 111(27):9983–8. Epub 2014/06/26. https://doi.org/10.1073/pnas.1314939111 PMID: 24961373.

32. Rusu V, Hoch E, Mercader JM, Tenen DE, Gymrek M, Hartigan CR, et al. Type 2 Diabetes Variants Disrupt Function of SLC16A11 through Two Distinct Mechanisms. Cell. 2017; 170(1):199–212 e20. Epub 2017/07/01. https://doi.org/10.1016/j.cell.2017.06.011 PMID: 28666119.

33. Masuo Y, Ohba Y, Yamada K, Al-Shammari AH, Seba N, Nakamichi N, et al. Combination Metabolomics Approach for Identifying Endogenous Substrates of Carnitine/Organic Cation Transporter OCTN1. Pharm Res. 2018; 35(11):224. Epub 2018/10/04. https://doi.org/10.1007/s11095-018-2507-1 PMID: 30280275.

34. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics. 2007; 3(3):211–21. Epub 2007/09/01. https://doi.org/10.1007/s11306-007-0082-2 PMID: 24039616.

35. Kim HI, Raffler J, Lu W, Lee JJ, Abbey D, Saleheen D, et al. Fine Mapping and Functional Analysis Reveal a Role of SLC22A1 in Acylcarnitine Transport. Am J Hum Genet. 2017; 101(4):489–502. Epub 2017/09/26. https://doi.org/10.1016/j.ajhg.2017.08.008 PMID: 28942964.

36. Price PJ, Gregory EA. Relationship between in vitro growth promotion and biophysical and biochemical properties of the serum supplement. In Vitro. 1982; 18(6):576–84. Epub 1982/06/01. PMID: 7118138.

37. Hagos Y, Stein D, Ugele B, Burckhardt G, Bahn A. Human renal organic anion transporter 4 operates as an asymmetric urate transporter. J Am Soc Nephrol. 2007; 18(2):430–9. Epub 2007/01/19. https://doi.org/10.1681/ASN.2006040415 PMID: 17229912.

38. Zou L, Stecula A, Gupta A, Prasad B, Chien HC, Yee SW, et al. Molecular Mechanisms for Species Differences in Organic Anion Transporter 1, OAT1: Implications for Renal Drug Toxicity. Mol Pharmacol. 2018; 94(1):689–99. Epub 2018/05/04. https://doi.org/10.1124/mol.117.111153 PMID: 29720497.

39. Rizwan AN, Krick W, Burckhardt G. The chloride dependence of the human organic anion transporter 1 (hOAT1) is blunted by mutation of a single amino acid. J Biol Chem. 2007; 282(18):13402–9. Epub 2007/03/14. https://doi.org/10.1074/jbc.M609849200 PMID: 17353191.

40. Sahlin S, Ahlberg J, Reihner E, Stahlberg D, Einarsson K. Cholesterol metabolism in human gallbladder mucosa: relationship to cholesterol gallstone disease and effects of chenodeoxycholic acid and ursodeoxycholic acid treatment. Hepatology. 1992; 16(2):320–6. Epub 1992/08/01. https://doi.org/10.1002/hep.1840160207 PMID: 1639340.

41. Sarenac TM, Mikov M. Bile Acid Synthesis: From Nature to the Chemical Modification and Synthesis and Their Applications as Drugs and Nutrients. Front Pharmacol. 2018; 9:939. Epub 2018/10/16. https://doi.org/10.3389/fphar.2018.00939 PMID: 30319399.

42. Choi MH, Kim KR, Chung BC. Simultaneous determination of urinary androgen glucuronides by high temperature gas chromatography-mass spectrometry with selected ion monitoring. Steroids. 2000; 65 (1):54–9. Epub 2000/01/07. https://doi.org/10.1016/s0039-128x(99)00082-3 PMID: 10624837.

43. Strahm E, Kohler I, Rudaz S, Martel S, Carrupt PA, Veuthey JL, et al. Isolation and quantification by high-performance liquid chromatography-ion-trap mass spectrometry of androgen sulfoconjugates in human urine. J Chromatogr A. 2008; 1196–1197:153–60. Epub 2008/05/27. https://doi.org/10.1016/j.chroma.2008.04.066 PMID: 18501914.

44. Wu W, Baker ME, Eraly SA, Bush KT, Nigam SK. Analysis of a large cluster of SLC22 transporter genes, including novel USTs, reveals species-specific amplification of subsets of family members. Physiol Genomics. 2009; 38(2):116–24. Epub 2009/05/07. https://doi.org/10.1152/physiolgenomics.90309.2008 PMID: 19417012.

45. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. Nucleic Acids Res. 2018; 46(D1):D754–D61. Epub 2017/11/21. https://doi.org/10.1093/nar/gkx1098 PMID: 29155950.

46. Ensembl Release 95. Gene: SLC22A24 Orthologues. 2019.

47. National Center for Biotechnology Information. Slc22a24 solute carrier family 22, member 24 [Rattus norvegicus (Norway rat)] 2018 [cited 2018]. https://www.ncbi.nlm.nih.gov/gene?term=(slc22a24[gene])%20AND%20(Rattus%20norvegicus[orgn])%20AND%20alive[prop]%20NOT%20newentry[gene]&sort=weight.

48. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, et al. An integrated map of structural variation in 2,504 human genomes. Nature. 2015; 526(7571):75–81. Epub 2015/10/04. https://doi.org/10.1038/nature15394 PMID: 26432246.

49. Wu H, Uchimura K, Donnelly EL, Kirita Y, Morris SA, Humphreys BD. Comparative Analysis and Refinement of Human PSC-Derived Kidney Organoid Differentiation with Single-Cell Transcriptomics. Cell Stem Cell. 2018; 23(6):869–81 e8. Epub 2018/11/20. https://doi.org/10.1016/j.stem.2018.10.010 PMID: 30449713.

50. Chang CY, Picotti P, Huttenhain R, Heinzelmann-Schwarz V, Jovanovic M, Aebersold R, et al. Protein significance analysis in selected reaction monitoring (SRM) measurements. Mol Cell Proteomics. 2012; 11(4):M111 014662. Epub 2011/12/23. https://doi.org/10.1074/mcp.M111.014662 PMID: 22190732.

51. Wu H, Malone AF, Donnelly EL, Kirita Y, Uchimura K, Ramakrishnan SM, et al. Single-Cell Transcriptomics of a Human Kidney Allograft Biopsy Specimen Defines a Diverse Inflammatory Response. J Am Soc Nephrol. 2018; 29(8):2069–80. Epub 2018/07/08. https://doi.org/10.1681/ASN.2018020125 PMID: 29980650.

52. Nigam SK, Bush KT, Martovetsky G, Ahn SY, Liu HC, Richard E, et al. The organic anion transporter (OAT) family: a systems biology perspective. Physiol Rev. 2015; 95(1):83–123. Epub 2014/12/30. https://doi.org/10.1152/physrev.00025.2013 PMID: 25540139.

53. Diotel N, Charlier TD, Lefebvre d'Hellencourt C, Couret D, Trudeau VL, Nicolau JC, et al. Steroid Transport, Local Synthesis, and Signaling within the Brain: Roles in Neurogenesis, Neuroprotection, and Sexual Behaviors. Front Neurosci. 2018; 12:84. Epub 2018/03/09. https://doi.org/10.3389/fnins.2018.00084 PMID: 29515356.

54. Hammes SR, Levin ER. Impact of estrogens in males and androgens in females. J Clin Invest. 2019; 129(5):1818–26. Epub 2019/05/02. https://doi.org/10.1172/JCI125755 PMID: 31042159.

55. Jin G, Sun J, Kim ST, Feng J, Wang Z, Tao S, et al. Genome-wide association study identifies a new locus JMJD1C at 10q21 that may influence serum androgen levels in men. Hum Mol Genet. 2012; 21 (23):5222–8. Epub 2012/09/01. https://doi.org/10.1093/hmg/dds361 PMID: 22936694.

56. Cropp CD, Komori T, Shima JE, Urban TJ, Yee SW, More SS, et al. Organic anion transporter 2 (SLC22A7) is a facilitative transporter of cGMP. Mol Pharmacol. 2008; 73(4):1151–8. Epub 2008/01/25. https://doi.org/10.1124/mol.107.043117 PMID: 18216183.

57. Enomoto A, Kimura H, Chairoungdua A, Shigeta Y, Jutabha P, Cha SH, et al. Molecular identification of a renal urate anion exchanger that regulates blood urate levels. Nature. 2002; 417(6887):447–52. Epub 2002/05/25. https://doi.org/10.1038/nature742 PMID: 12024214.

58. Skwara P, Schomig E, Grundemann D. A novel mode of operation of SLC22A11: Membrane insertion of estrone sulfate versus translocation of uric acid and glutamate. Biochem Pharmacol. 2017; 128:74–82. Epub 2016/12/29. https://doi.org/10.1016/j.bcp.2016.12.020 PMID: 28027879.

59. Sweet DH, Chan LM, Walden R, Yang XP, Miller DS, Pritchard JB. Organic anion transporter 3 (Slc22a8) is a dicarboxylate exchanger indirectly coupled to the Na+ gradient. Am J Physiol Renal Physiol. 2003; 284(4):F763–9. Epub 2002/12/19. https://doi.org/10.1152/ajprenal.00405.2002 PMID: 12488248.

60. Zhang X, Groves CE, Bahn A, Barendt WM, Prado MD, Rodiger M, et al. Relative contribution of OAT and OCT transporters to organic electrolyte transport in rabbit proximal tubule. Am J Physiol Renal Physiol. 2004; 287(5):F999–1010. Epub 2004/07/15. https://doi.org/10.1152/ajprenal.00156.2004 PMID: 15251863.

61. Anzai N, Jutabha P, Enomoto A, Yokoyama H, Nonoguchi H, Hirata T, et al. Functional characterization of rat organic anion transporter 5 (Slc22a19) at the apical membrane of renal proximal tubules. J Pharmacol Exp Ther. 2005; 315(2):534–44. Epub 2005/08/05. https://doi.org/10.1124/jpet.105.088583 PMID: 16079298.

62. Schomig E, Spitzenberger F, Engelhardt M, Martel F, Ording N, Grundemann D. Molecular cloning and characterization of two novel transport proteins from rat kidney. FEBS Lett. 1998; 425(1):79–86. Epub 1998/04/16. https://doi.org/10.1016/s0014-5793(98)00203-8 PMID: 9541011.

63. Yokoyama H, Anzai N, Ljubojevic M, Ohtsu N, Sakata T, Miyazaki H, et al. Functional and immunochemical characterization of a novel organic anion transporter Oat8 (Slc22a9) in rat renal collecting duct. Cell Physiol Biochem. 2008; 21(4):269–78. Epub 2008/04/29. https://doi.org/10.1159/000129385 PMID: 18441515.

64. Youngblood GL, Sweet DH. Identification and functional assessment of the novel murine organic anion transporter Oat5 (Slc22a19) expressed in kidney. Am J Physiol Renal Physiol. 2004; 287(2):F236–44. Epub 2004/04/08. https://doi.org/10.1152/ajprenal.00012.2004 PMID: 15068970.

65. Kwak JO, Kim HW, Oh KJ, Ko CB, Park H, Cha SH. Characterization of mouse organic anion transporter 5 as a renal steroid sulfate transporter. J Steroid Biochem Mol Biol. 2005; 97(4):369–75. Epub 2005/09/10. https://doi.org/10.1016/j.jsbmb.2005.06.028 PMID: 16150593.

66. Breljak D, Ljubojevic M, Balen D, Zlender V, Brzica H, Micek V, et al. Renal expression of organic anion transporter Oat5 in rats and mice exhibits the female-dominant sex differences. Histol Histopathol. 2010; 25(11):1385–402. Epub 2010/09/25. https://doi.org/10.14670/HH-25.1385 PMID: 20865662.

67. Tsuchida H, Anzai N, Shin HJ, Wempe MF, Jutabha P, Enomoto A, et al. Identification of a novel organic anion transporter mediating carnitine transport in mouse liver and kidney. Cell Physiol Biochem. 2010; 25(4–5):511–22. Epub 2010/03/25. https://doi.org/10.1159/000303060 PMID: 20332632.

68. Liang Q, Xu W, Hong Q, Xiao C, Yang L, Ma Z, et al. Rapid comparison of metabolites in humans and rats of different sexes using untargeted UPLC-TOFMS and an in-house software platform. Eur J Mass Spectrom (Chichester). 2015; 21(6):801–21. Epub 2016/01/15. https://doi.org/10.1255/ejms.1395 PMID: 26764310.

69. Guillemette C, Hum DW, Belanger A. Levels of plasma C19 steroids and 5 alpha-reduced C19 steroid glucuronides in primates, rodents, and domestic animals. Am J Physiol. 1996; 271(2 Pt 1):E348–53. Epub 1996/08/01. https://doi.org/10.1152/ajpendo.1996.271.2.E348 PMID: 8770030.

70. Uhlen M, Hallstrom BM, Lindskog C, Mardinoglu A, Ponten F, Nielsen J. Transcriptomics resources of human tissues and organs. Mol Syst Biol. 2016; 12(4):862. Epub 2016/04/06. https://doi.org/10.15252/msb.20155865 PMID: 27044256.

71. de Manuel M, Kuhlwilm M, Frandsen P, Sousa VC, Desai T, Prado-Martinez J, et al. Chimpanzee genomic diversity reveals ancient admixture with bonobos. Science. 2016; 354(6311):477–81. Epub 2016/10/30. https://doi.org/10.1126/science.aag2602 PMID: 27789843.

72. Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, et al. A draft sequence of the Neandertal genome. Science. 2010; 328(5979):710–22. Epub 2010/05/08. https://doi.org/10.1126/science.1188021 PMID: 20448178.

73. Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S, et al. A high-coverage genome sequence from an archaic Denisovan individual. Science. 2012; 338(6104):222–6. Epub 2012/09/01. https://doi.org/10.1126/science.1224344 PMID: 22936568.

74. Russel FG, Masereeuw R, van Aubel RA. Molecular aspects of renal anionic drug transport. Annu Rev Physiol. 2002; 64:563–94. Epub 2002/02/05. https://doi.org/10.1146/annurev.physiol.64.081501.155913 PMID: 11826280.

75. Sekine T, Miyazaki H, Endou H. Molecular physiology of renal organic anion transporters. Am J Physiol Renal Physiol. 2006; 290(2):F251–61. Epub 2006/01/13. https://doi.org/10.1152/ajprenal.00439.2004 PMID: 16403838.

76. Li CY, Basit A, Gupta A, Gaborik Z, Kis E, Prasad B. Major glucuronide metabolites of testosterone are primarily transported by MRP2 and MRP3 in human liver, intestine and kidney. J Steroid Biochem Mol Biol. 2019. Epub 2019/04/09. https://doi.org/10.1016/j.jsbmb.2019.03.027 PMID: 30959153.

77. Agopian AJ, Mitchell LE, Glessner J, Bhalla AD, Sewda A, Hakonarson H, et al. Genome-wide association study of maternal and inherited loci for conotruncal heart defects. PLoS One. 2014; 9(5):e96057. Epub 2014/05/08. https://doi.org/10.1371/journal.pone.0096057 PMID: 24800985.

78. Guo W, Bachman E, Li M, Roy CN, Blusztajn J, Wong S, et al. Testosterone administration inhibits hepcidin transcription and is associated with increased iron incorporation into red blood cells. Aging Cell. 2013; 12(2):280–91. Epub 2013/02/13. https://doi.org/10.1111/acel.12052 PMID: 23399021.

79. Pergialiotis V, Trakakis E, Parthenis C, Hatziagelaki E, Chrelias C, Thomakos N, et al. Correlation of platelet to lymphocyte and neutrophil to lymphocyte ratio with hormonal and metabolic parameters in women with PCOS. Horm Mol Biol Clin Investig. 2018; 34(3). Epub 2018/04/26. https://doi.org/10.1515/hmbci-2017-0073 PMID: 29694329.

80. Bachman E, Travison TG, Basaria S, Davda MN, Guo W, Li M, et al. Testosterone induces erythrocytosis via increased erythropoietin and suppressed hepcidin: evidence for a new erythropoietin/hemoglobin set point. J Gerontol A Biol Sci Med Sci. 2014; 69(6):725–35. Epub 2013/10/26. https://doi.org/10.1093/gerona/glt154 PMID: 24158761.

81. Nowak K, Jablonska E, Ratajczak-Wrona W. Neutrophils life under estrogenic and xenoestrogenic control. J Steroid Biochem Mol Biol. 2019; 186:203–11. Epub 2018/11/02. https://doi.org/10.1016/j.jsbmb.2018.10.015 PMID: 30381249.

82. Brochu M, Belanger A, Tremblay RR. Plasma levels of C-19 steroids and 5 alpha-reduced steroid glucuronides in hyperandrogenic and idiopathic hirsute women. Fertil Steril. 1987; 48(6):948–53. Epub 1987/12/01. https://doi.org/10.1016/s0015-0282(16)59589-2 PMID: 2960564.

83. Carmina E, Godwin AJ, Stanczyk FZ, Lippman JS, Lobo RA. The association of serum androsterone glucuronide with inflammatory lesions in women with adult acne. J Endocrinol Invest. 2002; 25(9):765–8. Epub 2002/10/26. https://doi.org/10.1007/BF03345509 PMID: 12398233.

84. Carmina E, Lobo RA. Evidence for increased androsterone metabolism in some normoandrogenic women with acne. J Clin Endocrinol Metab. 1993; 76(5):1111–4. Epub 1993/05/01. https://doi.org/10.1210/jcem.76.5.8496299 PMID: 8496299.

85. Carmina E, Stanczyk FZ, Matteri RK, Lobo RA. Serum androsterone conjugates differentiate between acne and hirsutism in hyperandrogenic women. Fertil Steril. 1991; 55(5):872–6. Epub 1991/05/01. PMID: 1827073.

86. Rocha M, Cardozo KHM, Carvalho VM, Bagatin E. ADT-G as a promising biomarker for peripheral hyperandrogenism in adult female acne. Dermatoendocrinol. 2017; 9(1):e1361571. Epub 2018/03/03. https://doi.org/10.1080/19381980.2017.1361571 PMID: 29497466.

87. Thompson DL, Horton N, Rittmaster RS. Androsterone glucuronide is a marker of adrenal hyperandrogenism in hirsute women. Clin Endocrinol (Oxf). 1990; 32(3):283–92. Epub 1990/03/01. https://doi.org/10.1111/j.1365-2265.1990.tb00868.x PMID: 2160872.

88. Accessible from Neale Lab. Imputed genotypes from HRC plus UK10K & 1000 Genomes reference panels as released by UK Biobank in March 2018 2018 [cited 2018 11/25/2018]. http://www.nealelab.is/uk-biobank/.

89. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am J Epidemiol. 1989; 129(4):687–702. Epub 1989/04/01. PMID: 2646917.

90. Yu B, Heiss G, Alexander D, Grams ME, Boerwinkle E. Associations Between the Serum Metabolome and All-Cause Mortality Among African Americans in the Atherosclerosis Risk in Communities (ARIC) Study. Am J Epidemiol. 2016; 183(7):650–6. Epub 2016/03/10. https://doi.org/10.1093/aje/kwv213 PMID: 26956554.

91. Yu B, Zheng Y, Alexander D, Morrison AC, Coresh J, Boerwinkle E. Genetic determinants influencing human serum metabolome among African Americans. PLoS Genet. 2014; 10(3):e1004212. Epub 2014/03/15. https://doi.org/10.1371/journal.pgen.1004212 PMID: 24625756.

92. Yu B, de Vries PS, Metcalf GA, Wang Z, Feofanova EV, Liu X, et al. Whole genome sequence analysis of serum amino acid levels. Genome Biol. 2016; 17(1):237. Epub 2016/11/26. https://doi.org/10.1186/s13059-016-1106-x PMID: 27884205.

93. Yu B, Li AH, Metcalf GA, Muzny DM, Morrison AC, White S, et al. Loss-of-function variants influence the human serum metabolome. Sci Adv. 2016; 2(8):e1600800. Epub 2016/09/08. https://doi.org/10.1126/sciadv.1600800 PMID: 27602404.

94. Stecula A, Schlessinger A, Giacomini KM, Sali A. Human Concentrative Nucleoside Transporter 3 (hCNT3, SLC28A3) Forms a Cyclic Homotrimer. Biochemistry. 2017; 56(27):3475–83. Epub 2017/07/01. https://doi.org/10.1021/acs.biochem.7b00339 PMID: 28661652.

95. Shima JE, Komori T, Taylor TR, Stryke D, Kawamoto M, Johns SJ, et al. Genetic variants of human organic anion transporter 4 demonstrate altered transport of endogenous substrates. Am J Physiol Renal Physiol. 2010; 299(4):F767–75. Epub 2010/07/30. https://doi.org/10.1152/ajprenal.00312.2010 PMID: 20668102.

96. Urban TJ, Gallagher RC, Brown C, Castro RA, Lagpacan LL, Brett CM, et al. Functional genetic diversity in the high-affinity carnitine transporter OCTN2 (SLC22A5). Mol Pharmacol. 2006; 70(5):1602–11. Epub 2006/08/26. https://doi.org/10.1124/mol.106.028126 PMID: 16931768.

97. Urban TJ, Brown C, Castro RA, Shah N, Mercer R, Huang Y, et al. Effects of genetic variation in the novel organic cation transporter, OCTN1, on the renal clearance of gabapentin. Clin Pharmacol Ther. 2008; 83(3):416–21. Epub 2007/07/05. https://doi.org/10.1038/sj.clpt.6100271 PMID: 17609685.

98. Yee SW, Nguyen AN, Brown C, Savic RM, Zhang Y, Castro RA, et al. Reduced renal clearance of cefotaxime in asians with a low-frequency polymorphism of OAT3 (SLC22A8). J Pharm Sci. 2013; 102(9):3451–7. Epub 2013/05/08. https://doi.org/10.1002/jps.23581 PMID: 23649425.

99. Chen J, Terada T, Ogasawara K, Katsura T, Inui K. Adaptive responses of renal organic anion transporter 3 (OAT3) during cholestasis. Am J Physiol Renal Physiol. 2008; 295(1):F247–52. Epub 2008/05/16. https://doi.org/10.1152/ajprenal.00139.2008 PMID: 18480179.

100. Chen EC, Khuri N, Liang X, Stecula A, Chien HC, Yee SW, et al. Discovery of Competitive and Noncompetitive Ligands of the Organic Cation Transporter 1 (OCT1; SLC22A1). J Med Chem. 2017; 60(7):2685–96. Epub 2017/02/24. https://doi.org/10.1021/acs.jmedchem.6b01317 PMID: 28230985.

101. Khuri N, Zur AA, Wittwer MB, Lin L, Yee SW, Sali A, et al. Computational Discovery and Experimental Validation of Inhibitors of the Human Intestinal Transporter OATP2B1. J Chem Inf Model. 2017; 57(6):1402–13. Epub 2017/06/01. https://doi.org/10.1021/acs.jcim.6b00720 PMID: 28562037.

102. Irannejad R, Pessino V, Mika D, Huang B, Wedegaertner PB, Conti M, et al. Functional selectivity of GPCR-directed drug action through location bias. Nat Chem Biol. 2017; 13(7):799–806. Epub 2017/05/30. https://doi.org/10.1038/nchembio.2389 PMID: 28553949.

103. Irannejad R, Tomshine JC, Tomshine JR, Chevalier M, Mahoney JP, Steyaert J, et al. Conformational biosensors reveal GPCR signalling from endosomes. Nature. 2013; 495(7442):534–8. Epub 2013/03/22. https://doi.org/10.1038/nature12000 PMID: 23515162.

104. Tian X, Irannejad R, Bowman SL, Du Y, Puthenveedu MA, von Zastrow M, et al. The alpha-Arrestin ARRDC3 Regulates the Endosomal Residence Time and Intracellular Signaling of the beta2-Adrenergic Receptor. J Biol Chem. 2016; 291(28):14510–25. Epub 2016/05/27. https://doi.org/10.1074/jbc.M116.716589 PMID: 27226565.

105. Dehaven CD, Evans AM, Dai H, Lawton KA. Organization of GC/MS and LC/MS metabolomics data into chemical libraries. J Cheminform. 2010; 2(1):9. Epub 2010/10/20. https://doi.org/10.1186/1758-2946-2-9 PMID: 20955607.

106. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004; 32(5):1792–7. Epub 2004/03/23. https://doi.org/10.1093/nar/gkh340 PMID: 15034147.

107. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees 2010. http://www.phylo.org/sub_sections/portal/sc2010_paper.pdf.

108. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014; 30(9):1312–3. Epub 2014/01/24. https://doi.org/10.1093/bioinformatics/btu033 PMID: 24451623.

109. Huson DH, Scornavacca C. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Syst Biol. 2012; 61(6):1061–7. Epub 2012/07/12. https://doi.org/10.1093/sysbio/sys062 PMID: 22780991.

110. Deng D, Sun P, Yan C, Ke M, Jiang X, Xiong L, et al. Molecular basis of ligand recognition and transport by glucose transporters. Nature. 2015; 526(7573):391–6. Epub 2015/07/16. https://doi.org/10.1038/nature14655 PMID: 26176916.

111. Pei J, Kim BH, Grishin NV. PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res. 2008; 36(7):2295–300. Epub 2008/02/22. https://doi.org/10.1093/nar/gkn072 PMID: 18287115.

112. Shen MY, Sali A. Statistical potential for assessment and prediction of protein structures. Protein Sci. 2006; 15(11):2507–24. Epub 2006/11/01. https://doi.org/10.1110/ps.062416606 PMID: 17075131.

113. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, et al. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucleic Acids Res. 2007; 35(Web Server issue):W522–5. Epub 2007/05/10. https://doi.org/10.1093/nar/gkm276 PMID: 17488841.

114. Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, et al. Improvements to the APBS biomolecular solvation software suite. Protein Sci. 2018; 27(1):112–28. Epub 2017/08/25. https://doi.org/10.1002/pro.3280 PMID: 28836357.

115. Nichols L, Freund M, Ng C, Kau A, Parisi M, Taylor A, et al. The National Institutes of Health Neurobiobank: a federated national network of human brain and tissue repositories. Biol Psychiatry. 2014; 75 (12):e21–2. Epub 2013/10/01. https://doi.org/10.1016/j.biopsych.2013.07.039 PMID: 24074636.

116. Cheung KWK, van Groen BD, Spaans E, van Borselen MD, De Brujin ACJM, Simons-Oosterhuis Y, et al. A comprehensive analysis of ontogeny of renal drug transporters: mRNA analyses, quantitative proteomics and localization. Clinical Pharmacology and Therapeutics submitted for publication.

117. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat Methods. 2009; 6(5):359–62. Epub 2009/04/21. https://doi.org/10.1038/nmeth.1322 PMID: 19377485.